



TUGAS AKHIR - SS141501

**PENGELOMPOKAN DESA DI KABUPATEN
BONDOWOSO BERDASARKAN DATA CAMPURAN
NUMERIK DAN KATEGORIK MENGGUNAKAN
METODE ENSEMBEL ROCK**

**RAHMAT RAMDHANY
NRP 1314 100 004**

**Dosen Pembimbing
Dr. Santi Puteri Rahayu, M.Si**

**PROGRAM STUDI SARJANA
DEPARTEMEN STATISTIKA
FAKULTAS MATEMATIKA, KOMPUTASI, DAN SAINS DATA
INSTITUT TEKNOLOGI SEPULUH NOPEMBER
SURABAYA 2018**



TUGAS AKHIR - SS141501

**PENGELOMPOKAN DESA
DI KABUPATEN BONDOWOSO BERDASARKAN
DATA CAMPURAN NUMERIK DAN KATEGORIK
MENGUNAKAN METODE ENSEMBEL ROCK**

**RAHMAT RAMDHANY
NRP 1314 100 004**

**Dosen Pembimbing
Dr. Santi Puteri Rahayu, M.Si**

**PROGRAM STUDI SARJANA
DEPARTEMEN STATISTIKA
FAKULTAS MATEMATIKA KOMPUTASI DAN SAINS DATA
INSTITUT TEKNOLOGI SEPULUH NOPEMBER
SURABAYA 2018**



FINAL PROJECT - SS 141501

**VILLAGE CLASSIFICATION
IN BONDOWOSO DISTRICT BASED ON
NUMERICAL AND CATEGORICAL MIXED DATA
USING ROCK ENSEMBLE METHOD**

**RAHMAT RAMDHANY
NRP 1314 100 004**

**Supervisor
Dr. Santi Puteri Rahayu, M.Si**

**UNDERGRADUATE PROGRAMME
DEPARTMENT OF STATISTICS
FACULTY OF MATHEMATICS, COMPUTING, AND DATA SCIENCES
INSTITUT TEKNOLOGI SEPULUH NOPEMBER
SURABAYA 2018**

LEMBAR PENGESAHAN
PENGELOMPOKAN DESA DI KABUPATEN
BONDOWOSO BERDASARKAN DATA CAMPURAN
NUMERIK DAN KATEGORIK MENGGUNAKAN
METODE ENSEMBEL ROCK

TUGAS AKHIR

Diajukan Untuk Memenuhi Salah Satu Syarat
Memperoleh Gelar Sarjana Sains
pada

Program Studi Sarjana Departemen Statistika
Fakultas Matematika, Komputasi, dan Sains Data
Institut Teknologi Sepuluh Nopember

Oleh :

Rahmat Ramdhany
NRP. 1314 100 004

Disetujui oleh Pembimbing :

Dr. Santi Puteri Rahayu, M.Si
NIP. 19750115 199903 2 003

()

Mengetahui,
Kepala Departemen


Dr. Suhartono
NIP. 19710929 199512 1 001

SURABAYA, 24 JANUARI 2018

PENGELOMPOKAN DESA DI KABUPATEN BONDOWOSO BERDASARKAN DATA CAMPURAN NUMERIK DAN KATEGORIK MENGGUNAKAN METODE ENSEMBEL ROCK

Nama : Rahmat Ramdhany
NRP : 1314 100 004
Departemen : Statistika
Dosen Pembimbing : Dr. Santi Puteri Rahayu, M.Si

Abstrak

Kabupaten Bondowoso merupakan salah satu kabupaten yang tergolong sebagai daerah tertinggal. Penentuan daerah tertinggal dilakukan oleh Bappenas berdasarkan Indeks Pembangunan Desa (IPD). Indikator dalam IPD tersusun atas data campuran yang berskala numerik dan kategorik. Untuk dapat mengatasi permasalahan daerah tertinggal perlu dilakukan penanganan mulai dari sub komponen terkecil di daerah, yaitu desa apa yang tergolong sebagai desa tertinggal. Salah satu metode statistika untuk dapat menentukan kelompok desa tertinggal adalah analisis pengelompokan atau analisis cluster. Namun karena variabel yang dihadapi tersusun atas data berskala campuran maka digunakan metode pengembangan dari analisis cluster yaitu analisis cluster dengan metode ensemble ROCK. Hasil pengelompokan desa di Kabupaten Bondowoso berdasarkan metode ensemble ROCK diperoleh 2 cluster dimana cluster 1 berisikan 71 desa dan cluster 2 berisikan 148 desa. Berdasarkan karakteristik dalam tiap cluster tersebut, cluster 1 menunjukkan kondisi yang paling rendah sehingga cluster 1 dianggap sebagai kelompok desa tertinggal.

Kata Kunci : Desa, Ensemble, Kategorik, Numerik, ROCK

(Halaman ini sengaja dikosongkan)

VILLAGE CLASSIFICATION IN BONDOWOSO DISTRICT BASED ON NUMERICAL AND CATEGORICAL MIXED DATA USING ROCK ENSEMBLE METHOD

Name : Rahmat Ramdhany
NRP : 1314 100 004
Department : Statistics
Supervisor : Dr. Santi Puteri Rahayu, M.Si

Abstract

Bondowoso is one of the districts that are categorized as underdeveloped areas. The underdeveloped areas determined by Bappenas based on the Village Development Index (IPD). Indicators in IPD are composed by numerical and categorical mixed data, to overcome the problem of underdeveloped areas, it is important to determine the villages which are classified as underdeveloped areas. One of statistical method to determine the group of underdeveloped villages is cluster analysis. However, because the variables encountered are composed of mixed-scale data, so the cluster analysis method used is ROCK ensemble methods. The result of village classification in Kabupaten Bondowoso based on ROCK ensemble method obtained 2 cluster where first cluster contains 71 villages and second cluster contains 148 villages. Based on the characteristics in each cluster, first cluster shows the lowest condition, so that the first cluster is considered as a underdeveloped villages group.

Keywords : Categorical, Ensemble, Numerical, ROCK, Villages

(halaman ini sengaja dikosongkan)

KATA PENGANTAR

Puji syukur kehadiran Allah SWT, Tuhan Yang Maha Esa. Berkat rahmat dan ridho-Nya penulis dapat menyelesaikan laporan Tugas Akhir yang berjudul **“Pengelompokan Desa Di Kabupaten Bondowoso Berdasarkan Data Campuran Numerik Dan Kategorik Menggunakan Metode Ensembl ROCK”** dengan lancar.

Keberhasilan penyusunan Tugas Akhir ini tidak lepas dari banyaknya bantuan dan dukungan yang diberikan dari berbagai pihak. Penulis mengucapkan terima kasih kepada:

1. Bapak Dr. Suhartono selaku Kepala Departemen Statistika dan Bapak Dr. Sutikno, M.Si selaku Koordinator Program Studi Sarjana yang telah memberikan fasilitas untuk kelancaran penyelesaian Tugas Akhir.
2. Ibu Dr. Santi Puteri Rahayu, M.Si selaku dosen pembimbing yang dengan sabar telah memberikan bimbingan, ilmu, saran, dan dukungan selama penyusunan Tugas Akhir.
3. Bapak Dr.rer.pol Heri Kuswanto, M.Si dan Bapak Dr. Ir. Setiawan, MS selaku dosen penguji yang telah memberikan banyak bantuan dan saran untuk Tugas Akhir ini.
4. Bapak Drs. Haryono, MSIE selaku dosen wali yang telah memberikan nasehat dan semangat selama masa perkuliahan.
5. Ibu dan Bapak tercinta, kakak tersayang serta keluarga yang selalu memberikan dukungan, kasih sayang, semangat, dan doa yang tidak pernah putus kepada penulis.

Penulis berharap hasil Tugas Akhir ini dapat bermanfaat bagi kita semua. Semoga kebaikan dan bantuan yang telah diberikan kepada penulis dibalas dengan kebaikan yang lebih besar lagi oleh Tuhan Yang Maha Esa. Aamiin.

Surabaya, Januari 2018

Penulis

(halaman ini sengaja dikosongkan)

DAFTAR ISI

HALAMAN JUDUL	i
LEMBAR PENGESAHAN	v
ABSTRAK	vii
ABSTRACT	ix
KATA PENGANTAR	xi
DAFTAR ISI	xiii
DAFTAR GAMBAR	xvii
DAFTAR TABEL	xix
DAFTAR LAMPIRAN	xxi
BAB I PENDAHULUAN	1
1.1 Latar Belakang	1
1.2 Rumusan Masalah	4
1.3 Tujuan Penelitian	5
1.4 Manfaat Penelitian	5
1.5 Batasan Masalah	6
BAB II TINJAUAN PUSTAKA	7
2.1 Statistika Deskriptif	7
2.2 Analisis Kelompok (<i>Cluster Analysis</i>).....	8
2.3 Analisis Faktor	11
2.4 Metode Pengelemokan	11
2.4.1 Pengelompokan Data Numerik	12
2.4.2 Pengelompokan Data Kategorik	15
2.5 Pengelompokan Data Campuran Numerik dan Kategorik	17

2.6 Kinerja Hasil Pengelompokan.....	19
2.7 <i>Two Step Cluster</i>	21
2.8 <i>One Way</i> ANOVA.....	22
2.9 <i>One Way</i> MANOVA.....	23
2.10 Uji Mann Whitney	25
2.11 Desa Tertinggal	25
2.12 Gambaran Umum Kabupaten Bondowoso.....	30
BAB III METODOLOGI PENELITIAN	33
3.1 Sumber Data 33	
3.2 Variabel Penelitian	33
3.3 Langkah Penelitian.....	41
BAB IV ANALISIS DAN PEMBAHASAN.....	47
4.1 Karakteristik Desa di Kabupaten Bondowoso	47
4.2 Analisis Faktor pada Data Numerik	50
4.3 Pengelompokan Data Numerik	52
4.4 Pengelompokan Data Kategorik	54
4.5 Pengelompokan Data Campuran Menggunakan Metode Ensemble ROCK.....	55
4.6 Pengelompokan dengan Metode <i>Two Step Cluster</i>	56
4.7 Perbandingan Metode Ensemble ROCK dengan <i>Two Step Cluster</i>	58
4.8 Karakteristik Desa Hasil Pengelompokan Ensemble ROCK.....	58
4.9 Pengujian Beda Rata-rata Antar Kelompok	66
BAB V KESIMPULAN DAN SARAN	71
5.1 Kesimpulan	71

5.2 Saran	73
DAFTAR PUSTAKA	75
LAMPIRAN	77

(Halaman ini sengaja dikosongkan)

DAFTAR GAMBAR

Gambar 2.1	Peta Kabupaten Bondowoso	31
Gambar 3.1	Diagram Alir Penelitian.....	44
Gambar 4.1	Persebaran Desa hasil Final Cluster.....	59
Gambar 4.2	Perbandingan Karakteristik Sarana Kesehatan, (a) Poskesdes, (b) Tempat Praktik Bidan, dan (c) Tempat Praktik Dokter.	61
Gambar 4.3	Perbandingan Karakteristik (a) Pasar Permanen dan (b) Toko Kelontong.....	62
Gambar 4.4	Perbandingan Karakteristik Persentase Keluarga Pengguna Listrik PLN.....	62
Gambar 4.5	Perbandingan Karakteristik Keuangan Desa, (a) PAD dan (b) ADD	63
Gambar 4.6	Perbandingan Karakteristik Sarana Pendidikan, (a) SD/MI, (b) SMP/MTS, dan (c) Jarak Menuju SMA/SMK/MA terdekat	64

(halaman ini sengaja dikosongkan)

DAFTAR TABEL

Tabel 2.1 Tabel Kontingensi Data Biner	9
Tabel 2.2 Ukuran Jarak Data Biner	9
Tabel 2.3 Tabel Transformasi Nilai Wilk's Lambda	23
Tabel 3.1 Variabel Penelitian.....	33
Tabel 4.1 Karakteristik Data Numerik.....	47
Tabel 4.2 Karakteristik Data Kategorik	49
Tabel 4.3 Eigen Value dan Persentase Variansi Kumulatif	51
Tabel 4.4 Nilai Loading Factor.....	51
Tabel 4.5 Nilai Pseudo-F Pengelompokan Data Numerik.....	53
Tabel 4.6 Rasio S_W dan S_B dalam Pengelompokan ROCK	54
Tabel 4.7 Struktur Data dengan Variabel Baru.....	55
Tabel 4.8 Rasio S_W dan S_B Pengelompokan Ensembel ROCK ..	56
Tabel 4.9 Hasil Pengelompokan Numerik dan Kategorik Two Step Cluster.....	57
Tabel 4.10 Final Cluster Metode Two Step Cluster	57
Tabel 4.11 Perbandingan Rasio Ensembel ROCK dengan Two Step Cluster	58
Tabel 4.12 Perbandingan Karakteristik Numerik Antar Cluster.	59
Tabel 4.13 Perbandingan Karakteristik Variabel Kategorik	64
Tabel 4.14 Hasil Pengujian ANOVA One Way	67
Tabel 4.15 Hasil Pengujian Mann Whittney	68

(halaman ini sengaja dikosongkan)

DAFTAR LAMPIRAN

Lampiran 1	Data Penelitian Variabel Numerik.....	77
Lampiran 2	Data Penelitian Variabel Kategorik.....	78
Lampiran 3	Syntax Pengelompokan Data Numerik.....	79
Lampiran 4	Syntax Pengelompokan Data Kategorik.....	81
Lampiran 5	Syntax Pengelompokan Ensembel ROCK.....	87
Lampiran 6	Output Statistika Deskriptif Seluruh Data.....	89
Lampiran 7	Pemeriksaan KMO dan Uji Bartlet.....	90
Lampiran 8	Persentase Variansi Kumulatif.....	91
Lampiran 9	<i>Rotated Component Matrix</i>	92
Lampiran 10	Output Pengelompokan Data Numerik	93
Lampiran 11	Output Pengelompokan Data Kategorik.....	97
Lampiran 12	Output Pengelompokan Ensembel ROCK.....	101
Lampiran 13	Output Pengelompokan Two Step Cluster.....	105
Lampiran 14	Syntax dan Output Mshapiro Test.....	109
Lampiran 15	Output Pengujian Beda Mean MANOVA.....	110
Lampiran 16	Output Pengujian ANOVA.....	111
Lampiran 17	Output Uji Mann-Whitney.....	113
Lampiran 18	Surat Keterangan Data dari BPS.....	117

(halaman ini sengaja dikosongkan)

BAB I

PENDAHULUAN

1.1 Latar Belakang

Daerah tertinggal merupakan suatu daerah dengan kabupaten yang masyarakat dan wilayahnya relatif kurang berkembang dibandingkan daerah lain dalam skala nasional. Pada tahun 2004 terdapat 199 kabupaten tertinggal, kemudian pada tahun 2009 menurun menjadi 183 kabupaten dan pada tahun 2014 tersisa 122 kabupaten yang masih tergolong sebagai daerah tertinggal. Jumlah kabupaten tertinggal di Kawasan Indonseia Timur mencapai 103 kabupaten atau 84,42 persen dari total 122 kabupaten daerah tertinggal, sedangkan sisanya sebanyak 19 kabupaten tertinggal atau 15,57 persen berada di Kawasan Barat Indonesia (Kementerian DPDTT, 2016).

Melihat dalam lingkup yang lebih kecil yaitu desa, menurut Kementerian DPDTT sebanyak 39.901 desa dari 73.709 desa yang ada di Indonesia dikategorikan sebagai desa tertinggal. Ini artinya lebih dari 50 persen desa yang terdapat di Indonesia masih tergolong sebagai desa tertinggal. Dalam Permendagri no 2 tahun 2017 tentang Standar Pelayanan Minimal Desa dijelaskan bahwa desa merupakan kesatuan masyarakat hukum yang memiliki batas wilayah yang berwenang untuk mengatur dan mengurus urusan pemerintahan, kepentingan masyarakat setempat berdasarkan prakarsa masyarakat, hak asal usul, dan/atau hak tradisional yang diakui dan dihormati dalam sistem pemerintahan Negara Kesatuan Republik Indonesia. Dengan demikian unsur desa sebagai kesatuan masyarakat hukum terkecil di Indonesia mampu menjadi pemutus akar permasalahan dari tingginya kasus daerah tertinggal di Indonesia. Diharapkan dengan kebijakan yang tepat untuk menangani masalah desa tertinggal sehingga akan memberikan kemajuan pembangunan untuk Indonesia.

Untuk dapat menetapkan kebijakan yang tepat dalam pembangunan desa tertinggal, diperlukan pemetaan terhadap profil desa berdasarkan tingkat perkembangannya. Pemetaan kelompok-

kelompok desa dapat dihasilkan melalui salah satu metode statistika yaitu analisis pengelompokan (*Cluster Analysis*). Pembangunan desa merupakan konsep multidimensional yang kompleks, sehingga diperlukan dimensi dan indikator yang mampu menggambarkan secara nyata tingkat kemajuan desa yang dipotret pada suatu waktu.

Dalam penentuan desa tertinggal terdapat lima aspek atau dimensi yang menjadi pertimbangan. Pertama adalah aspek pelayanan dasar yang mewakili aspek pelayanan dasar untuk mewujudkan bagian dari kebutuhan dasar khususnya untuk pendidikan dan kesehatan. Kedua adalah aspek kebutuhan infrastruktur yang mewakili kebutuhan dasar akan sarana, prasarana, pengembangan ekonomi lokal dan pemanfaatan sumber daya alam secara berkelanjutan. Ketiga aspek aksesibilitas dan transformasi yang mana dipisahkan sebagai indikator tersendiri dalam pembangunan desa dengan mempertimbangkan sarana dan prasarana transformasi memiliki kekhususan dan prioritas pembangunan desa sebagai penghubung kegiatan sosial ekonomi di desa. Keempat adalah aspek pelayanan umum yang merupakan upaya pemenuhan kebutuhan pelayanan atas barang, jasa dan/atau pelayanan administratif. Serta yang kelima adalah aspek penyelenggaraan pemerintahan yang mewakili kinerja pemerintahan desa merupakan bentuk pelayanan administratif yang disenggarakan penyelenggara pelayanan bagi warga yang dalam hal ini adalah pemerintah. Umumnya indikator-indikator ini tidak hanya berskala data numerik, melainkan juga berskala kategorik. Dengan kasus seperti ini analisis pengelompokan konvensional kurang mampu memberikan hasil pengelompokan yang cukup baik, sehingga perlu pengembangan terhadap metode analisis pengelompokan konvensional yang mampu mengatasi masalah pada data berskala campuran numerik dan kategorik.

Sebagaimana yang tertuang dalam Peraturan Presiden nomor 131 tahun 2015 tentang Penetapan Daerah Tertinggal tahun 2015-2019 terdapat 122 daerah tertinggal dimana empat kabupaten diantaranya berada di Provinsi Jawa Timur. Kabupaten yang

tergolong sebagai daerah tertinggal di Provinsi Jawa Timur adalah Kabupaten Sampang, Bangkalan, Situbondo dan Bondowoso. Kabupaten Bondowoso memiliki persentase desa tertinggal sebesar 8.13 persen dari total 219 desa/kelurahan yang berada di kabupaten tersebut (Bappenas & BPS, 2015). Tingginya persentase daerah tertinggal di Kabupaten Bondowoso cukup disayangkan karena potensi daerah yang dimilikinya cukup besar.

Dengan menggunakan analisis pengelompokan dapat dibuat profil desa-desa di Kabupaten Bondowoso menurut tingkat perkembangannya. Dalam penelitian ini digunakan metode pengembangan dari analisis pengelompokan dengan melibatkan data yang memiliki variabel campuran berskala numerik dan kategorik yaitu analisis pengelompokan menggunakan metode ensemble ROCK.

Metode yang seringkali dilakukan untuk pengelompokan data berskala campuran adalah dengan mentransformasi data kategorik menjadi data numerik dan sebaliknya. Dewangan, Sharma, dan Akasapu (2010) melakukan transformasi variabel kategorik ke dalam bentuk numerik, kemudian pengelompokan objek dilakukan dengan metode pengelompokan data numerik. Kelebihan metode transformasi adalah dapat mengurangi kompleksitas dalam komputasi. Akan tetapi, metode tersebut memiliki kelemahan dalam menentukan transformasi yang tepat agar tidak kehilangan banyak informasi dari original datanya. Terdapat juga metode pengelompokan ensemble untuk data campuran yang dikembangkan oleh He, Xu dan Deng pada tahun 2005. Pengelompokan ensemble adalah teknik pengelompokan untuk menggabungkan hasil pengelompokan beberapa algoritma untuk mendapatkan kelompok yang lebih baik (He, Xu, & Deng, A Cluster Ensemble Method For Clustering Categorical Data, 2005a). Selanjutnya He, Xu dan Deng kembali melakukan penelitian untuk pengelompokan persetujuan pengajuan kartu kredit dengan *Cluster Ensemble Based Mixed Data Clustering* (CEBMDC) yang menggunakan algoritma *Squeezer* untuk menggabungkan hasil pengelompokan masing-masing jenis data.

Didapatkan hasil bahwa algoritma *Squeezer* merupakan metode pengelompokan yang efektif digunakan untuk data yang berjumlah besar, sedangkan untuk data dengan jumlah yang relatif sedikit dapat digunakan metode ROCK.

Metode ROCK dikembangkan oleh Guha, Rastogi, dan Shim pada tahun 2000. Metode ROCK menggunakan konsep *link* untuk mengukur kesamaan/kedekatan antar pengamatan. Pada saat itu, Guha melakukan pengelompokan dengan metode ROCK untuk memisahkan Demokrat dari Partai Republik dengan data *Cogressional* Amerika Serikat dimana didapatkan hasil bahwa metode ROCK memiliki akurasi lebih baik daripada metode hirarki dengan sifat skalabilitas yang baik (Guha, Rastogi, & Shim, 2000). Dalam penelitian lainnya. Alvionita (2017) melakukan perbandingan kinerja antara metode ensemble ROCK dan SWFM pada studi kasus pengelompokan aksesori jeruk hasil fusi protoplasma. Hasil dari penelitian tersebut menunjukkan metode ROCK memberikan hasil pengelompokan yang lebih baik karena memiliki rasio perbandingan antara simpangan baku di dalam kelompok dengan simpangan baku antar kelompok yang lebih kecil daripada metode SWFM.

Oleh karena itu, dalam penelitian ini menggunakan metode pengelompokan ensemble ROCK ,dimana diharapkan akan mendapatkan hasil pengelompokan terbaik terhadap desa-desa di Kabupaten Bondowoso berdasarkan tingkat perkembangannya. Sehingga dari hasil pengelompokan tersebut didapatkan profil desa-desa di Kabupaten Bondowoso berdasarkan tingkat perkembangannya yang masih tergolong sebagai daerah tertinggal untuk selanjutnya berdasarkan pengelompokan tersebut dapat memberikan rekomendasi bagi pemerintah dalam menentukan kebijakan pembangunan daerah tertinggal.

1.2 Rumusan Masalah

Berdasarkan latar belakang diatas, permasalahan Kabupaten Bondowoso sebagai daerah tertinggal perlu segera dituntaskan melalui strategi pembangunan daerah yang didasarkan atas karakteristik dan kekhasan daerahnya mulai dari lingkup terkecil

yaitu desa. Oleh karena itu, rumusan masalah yang diangkat dalam penelitian ini sebagai berikut

1. Bagaimana karakteristik desa di Kabupaten Bondowoso berdasarkan indikator indikator pembangunan desa?
2. Bagaimana pengelompokan desa di Kabupaten Bondowoso dengan menggunakan metode Ensembl ROCK?
3. Bagaimana pengelompokan desa di Kabupaten Bondowoso dengan menggunakan metode *two step cluster*?
4. Bagaimana evaluasi kinerja kebaikan pengelompokan desa menggunakan metode ensembl ROCK dan *two step cluster*?

1.3 Tujuan Penelitian

Berdasarkan permasalahan yang telah diuraikan sebelumnya, maka tujuan dari penelitian ini adalah sebagai berikut

1. Untuk dapat menentukan karakteristik desa di Kabupaten Bondowoso berdasarkan indikator-indikator pembangunan desa.
2. Untuk dapat menentukan kelompok desa di Kabupaten Bondowoso dengan menggunakan metode ensembl ROCK.
3. Untuk dapat menentukan kelompok desa di Kabupaten Bondowoso dengan menggunakan metode *two step cluster*.
4. Untuk dapat menentukan metode terbaik dalam melakukan pengelompokan desa di Kabupaten Bondowoso.

1.4 Manfaat Penelitian

Hasil dari penelitian ini diharapkan mampu memberikan hasil pengelompokan desa-desa di Kabupaten Bondowoso berdasarkan indikator-indikator pembangunan desa yang terdiri dari data berskala numerik dan kategorikal. Setelah mendapatkan informasi mengenai kelompok desa-desa di Kabupaten Bondowoso yang tergolong sebagai daerah tertinggal, maka diharapkan hasil penelitian ini mampu menjadi rekomendasi bagi pemerintah dalam menyusun kebijakan sehingga dapat menyelesaikan persoalan daerah tertinggal tersebut. Selain itu, bagi peneliti dapat mengaplikasikan ilmu statistika dalam kehidupan

sehari-hari terutama dalam melakukan analisis pengelompokan terhadap data multivariat yang memiliki skala campuran

1.5 Batasan Masalah

Batasan masalah dalam penelitian ini adalah dalam melakukan pengujian perbedaan rata-rata antar kelompok hasil *final* cluster, secara multivariat pengujian dilakukan dengan menggunakan metode MANOVA. Oleh karena itu, dalam menguji perbedaan antar cluster hanya memperhatikan variabel yang berskala numerik. Namun untuk mengetahui variabel apa yang berbeda di antara cluster 1 dan cluster 2, variabel-variabel dalam penelitian ini dipandang sebagai variabel univariat, dimana untuk variabel numerik dilakukan pengujian menggunakan ANOVA *one way* dan untuk variabel kategorik dilakukan pengujian menggunakan uji Mann Whitney. Selain itu data numerik dianggap telah berdistribusi normal multivariat sebagai syarat dilakukannya pengujian beda rata-rata antar kelompok.

BAB II

TINJAUAN PUSTAKA

Pada bab ini akan dijelaskan teori-teori yang digunakan dalam melakukan analisis pengelompokan terhadap data multivariat yang berskala campuran numerik dan kategorikal. Data numerik merupakan data dengan variabel kuantitatif yang memberikan informasi numerik sedangkan data kategorik adalah data dengan variabel kualitatif yang dihasilkan dari penggolongan atau pengklasifikasian suatu data (Anderson & Sclove, 1974). Analisis multivariat adalah semua teknik statistik yang digunakan untuk menganalisis secara simultan beberapa pengukuran terhadap individu atau objek yang sedang diselidiki (Hair, Black, Babin, & Anderson, 2009). Analisis pengelompokan terhadap data multivariat yang berskala campuran tidak dapat dilakukan hanya secara parsial sehingga membutuhkan metode khusus untuk mampu menanganinya secara simultan. Dalam penelitian ini akan digunakan metode ensembel ROCK untuk mendapatkan hasil terbaik dalam pengelompokan data multivariat berskala campuran pada kasus pengelompokan desa tertinggal di Kabupaten Bondowoso. Dalam analisisnya pengelompokan terhadap data numerik menggunakan metode hirarki *agglomerative* sedangkan pengelompokan terhadap data kategorik menggunakan metode ROCK.

2.1 Statistika Deskriptif

Statistika deskriptif adalah metode-metode yang berkaitan dengan pengumpulan dan penyajian suatu gugus data sehingga memberikan informasi yang berguna (Walpole, 1993). Statistika deskriptif dapat digunakan untuk menjelaskan karakteristik suatu data. Penjelasan tersebut didapatkan dari pendefinisian ukuran-ukuran numerik yang dihitung dari pusat data tersebut. Untuk variabel berskala numerik digunakan ukuran pemusatan data *mean* atau rata-rata. Sedangkan untuk variabel berskala kategorik menggunakan ukuran pemusatan data modus yaitu nilai yang sering muncul pada suatu data. Nilai modus pada data tunggal

dapat di cari dengan melihat frekuensi dari data yang ada. Data dengan frekuensi terbanyak menunjukkan modus.

Selain itu, statistika deskriptif dapat juga disampaikan secara visual untuk memudahkan dalam pemberian informasi mengenai data yang dimiliki. Salah satu metodenya adalah dengan menggunakan *box plot*. *Box plot* merupakan salah satu cara dalam statistik deskriptif untuk menggambarkan secara grafik dari data numeris melalui lima ukuran yaitu nilai observasi terendah, quartil pertama, median, ruartil ketiga serta nilai observasi tertinggi sehingga *box plot* dapat juga digunakan untuk identifikasi data *outlier*.

2.2 Analisis Kelompok (*Cluster Analysis*)

Analisis kelompok (*cluster analysis*) merupakan salah satu teknik dalam analisis multivariat metode independen yang mereduksi (mengelompokkan) data. Analisis ini mempunyai tujuan utama untuk mengelompokkan objek-objek pengamatan menjadi beberapa kelompok berdasarkan karakteristik yang dimilikinya. Analisis kelompok mengelompokkan objek-objek sehingga setiap objek yang paling dekat kesamaannya dengan objek lain berada dalam kelompok yang sama, serta mempunyai kemiripan satu dengan yang lain (Johnson & Wicherin, 2007). Hasil analisis kelompok dipengaruhi oleh beberapa hal, antara lain yaitu objek yang dikelompokkan, variabel yang diamati, ukuran kemiripan dan ketidakmiripan, serta metode pengelompokan yang digunakan.

Ukuran kemiripan dan ketidakmiripan merupakan hal yang sangat mendasar dalam kelompok analisis. Algoritma pengelompokan menggunakan ukuran kemiripan atau ketidakmiripan digunakan untuk menggabungkan atau memisahkan objek dari suatu data. Ukuran kemiripan biasanya digunakan dalam algoritma pengelompokan untuk menganalisis data kategori, sedangkan ukuran ketidakmiripan digunakan oleh algoritma pengelompokan untuk menganalisis data numerik. Ukuran kemiripan dan ketidakmiripan pada umumnya diukur berdasarkan jarak. Salah satu faktor yang sangat berpengaruh

terhadap hasil dari kelompok yang dibentuk adalah jarak antara objek pengamatan (Sharma, 1996). Berikut ini merupakan metode-metode pengukuran jarak antara objek ke- i (x_i) dengan objek ke- j (x_j) berdasarkan karakteristik variabel yang dikelompokkan.

- a. Metode pengukuran jarak untuk variabel kategorik biner
Bila variabel yang diamati berupa variabel kategori biner yang hanya memiliki dua macam karakter yang berbeda (0,1) maka variabel yang diamati dapat dibentuk suatu tabel kontingensi seperti yang ditunjukkan pada Tabel 2.1. Perhitungan ukuran jarak antara variabel x_i dan x_j untuk pengukuran data biner dapat menggunakan beberapa ukuran yang disajikan pada Tabel 2.2

Tabel 2.1 Tabel Kontingensi Data Biner

Kategori x_i	Kategori x_j		Total
	1	0	
1	a	b	$a + b$
0	c	d	$c + d$
Total	$a + c$	$b + d$	$a + b + c + d$

Tabel 2.2 Ukuran Jarak Data Biner

Jenis	Rumus
<i>Russel and Rao</i>	$RR(x_i, x_j) = \frac{a}{a + b + c + d}$
<i>Simple Matching</i>	$SM(x_i, x_j) = \frac{a + d}{a + b + c + d}$
<i>Jaccard</i>	$JACCARD(x_i, x_j) = \frac{a}{a + b + c}$
<i>Dice Czekanowski</i>	$DICE(x_i, x_j) = \frac{2a}{2a + b + c}$

- b. Metode pengukuran jarak untuk variabel kategorik nominal
Pengamatan dengan variabel nominal maka pengukuran memiliki konsep yang sama dengan *simple matching coefficient* maupun *dice*, dimana kategorinya dapat lebih dari dua macam. Dengan jumlah variabel sebanyak m , maka rumus untuk pengukuran jarak variabel nominal antara x_i dan x_j dapat ditunjukkan pada persamaan 2.1

$$sim(x_i, x_j) = \frac{1}{m} \sum_{i=1}^m S_{ijl} \quad (2.1)$$

dimana $S_{ijl} = 1$ jika $x_{il} = x_{jl}$ dan $S_{ijl} = 0$ jika $x_{il} \neq x_{jl}$

- c. Metode pengukuran jarak untuk variabel kategori ordinal
Pengamatan dengan variabel ordinal konsep pengukuran jarak yang digunakan sama dengan metode untuk data numerik, dimana kategorinya dinyatakan sebagai suatu bilangan bulat. Salah satu metode yang dapat digunakan untuk variabel ordinal adalah jarak *manhattan*. Dengan jumlah variabel sebanyak m , maka rumus untuk pengukuran jarak variabel ordinal antara x_i dan x_j dapat ditunjukkan pada persamaan 2.2

$$sim(x_i, x_j) = \sum_{i=1}^m |x_{il} - x_{jl}| \quad (2.2)$$

- d. Metode pengukuran jarak untuk variabel numerik
Pengamatan yang memiliki jenis data numerik maka jarak yang dapat digunakan adalah jarak *euclidean*. Misalkan terdapat dua observasi dengan variabel-variabel berdimensi m yaitu $x_i = [x_1, x_2, \dots, x_m]^T$ dan $x_j = [x_1, x_2, \dots, x_m]^T$. Rumus untuk menghitung jarak *euclidean* dapat ditunjukkan pada persamaan 2.4

$$d_{ij} = \sqrt{(x_i - x_j)^T (x_i - x_j)} \quad (2.3)$$

Dengan dilakukannya analisis kelompok, maka diharapkan mendapatkan hasil pengelompokan yang baik. Hasil pengelompokan yang baik artinya dalam suatu kelompok mempunyai homogenitas yang tinggi antara anggota dalam suatu kelompok serta memiliki heterogenitas yang tinggi antara kelompok yang satu dengan kelompok lainnya (Hair, Black, Babin, & Anderson, 2009).

2.3 Analisis Faktor

Tujuan dari analisis faktor adalah untuk menggambarkan hubungan-hubungan kovarian antara beberapa variabel yang mendasari tetapi tidak teramati, kuantitas random yang disebut faktor. Vektor acak teramati X dengan p komponen, memiliki rata-rata dan matrik kovarian Σ . Model analisis faktor adalah sebagai berikut :

$$X_1 - \mu_1 = \ell_{11}F_1 + \ell_{12}F_2 + \dots + \ell_{1m}F_m + \varepsilon_1 \quad (2.4)$$

$$X_p - \mu_p = \ell_{p1}F_1 + \ell_{p2}F_2 + \dots + \ell_{pm}F_m + \varepsilon_p \quad (2.5)$$

Atau dapat ditulis dalam notasi matrik sebagai berikut

$$X_{p \times 1} = \mu_{(p \times 1)} + L_{(p \times m)}F_{(m \times 1)} + \varepsilon_{p \times 1} \quad (2.6)$$

μ = rata-rata variabel i

ε_i = faktor spesifik ke- i

F_j = *common factor* ke- j

ℓ_{ij} = *loading* dari variabel ke- i pada faktor ke- j

Bagian dari varian variabel ke- i dari m *common factor* disebut komunalitas ke- i yang merupakan jumlah kuadrat dari *loading* variabel ke- i pada m *common factor* dengan rumus

$$h_i^2 = \ell_{i1}^2 + \ell_{i2}^2 + \dots + \ell_{im}^2 \quad (2.7)$$

2.4 Metode Pengelemokan

Tahap pengelompokan dalam analisis kelompok dibedakan menurut jenis data yang dimiliki. Pada umumnya analisis kelompok terfokus pada data numerik, namun sering ditemukan kasus dengan data kategorik bahkan kasus dengan campuran data

numerik dan kategorik. Oleh karena itu dalam pengelompokannya menggunakan metode yang berbeda.

2.4.1 Pengelompokan Data Numerik

Pengelompokan data numerik dilakukan berdasarkan ukuran ketidakmiripan atau jarak untuk data numerik. Hasil pengelompokan disajikan dalam bentuk dendrogram (diagram pohon) yang memungkinkan penelusuran objek-objek yang diamati menjadi lebih mudah dan informatif. Teknik yang digunakan untuk pengelompokan meliputi metode hirarki dan metode non hirarki.

Pengelompokan hirarki dimulai dengan dua atau lebih objek yang mempunyai kesamaan paling dekat, kemudian proses diteruskan ke objek lain yang mempunyai kedekatan kedua. Analisis dilakukan hingga kelompok membentuk hirarki (tingkatan) yang jelas antar objek dari yang paling mirip sampai paling tidak mirip. Terdapat dua teknik pengelompokan dalam analisis kelompok hirarki yaitu teknik pembagian (*divisive*) dan teknik penggabungan (*agglomerative*).

Dalam teknik pengelompokan hirarki *agglomerative* setiap objek merupakan satu kelompok tersendiri. Lalu dua kelompok terdekat digabungkan dan seterusnya sehingga diperoleh satu kelompok yang berunsurkan semua objek. Untuk menggabungkan dua kelompok pada tahap awal, dimana tiap kelompok hanya terdiri dari satu objek, diperlukan ukuran ketidakmiripan antar objek. Bila suatu kelompok merupakan penggabungan dari beberapa kelompok sebelumnya, maka diperlukan ukuran ketidakmiripan antar kelompok, kelompok-kelompok dengan ukuran ketidakmiripan terkecil digabungkan menjadi kelompok yang baru.

Andaikan d_{uv} merupakan ukuran ketidakmiripan antara kelompok ke- u dengan kelompok ke- v dan $d_{w(uv)}$ merupakan ukuran ketidakmiripan antarak kelompok ke- w dengan kelompok (u,v) yang merupakan penggabungan antara kelompok ke- u dan

kelompok ke- v , maka beberapa teknik pengelompokan antar kelompok dinyatakan sebagai berikut (Johnson & Wicherin, 2007)

a. Pautan Tunggal (*Single Linkage/Nearest Neighbor*)

Prosedur ini didasarkan pada jarak terkecil atau jarak terdekat. Jika dua objek terpisah oleh jarak yang pendek maka kedua objek tersebut digabung menjadi satu kelompok dan demikian seterusnya. Ukuran jarak yang digunakan adalah

$$d_{w(uv)} = \min(d_{wu}, d_{wv}) \quad (2.8)$$

b. Pautan Lengkap (*Complete Linkage/Farthest Neighbor*)

Prosedur ini berlawanan dengan *single linkage* karena dalam prosedur ini didasarkan pada jarak terbesar atau jarak terjauh. Ukuran jarak yang digunakan adalah

$$d_{w(uv)} = \max(d_{wu}, d_{wv}) \quad (2.9)$$

c. Pautan Rataan (*Average Linkage Between Method/BAVERAGE*)

Prosedur dalam pautan rata-rata hampir sama dengan *single linkage* maupun *complete linkage*, namun kriteria yang digunakan adalah rata-rata jarak seluruh individu dalam suatu kelompok dengan jarak seluruh individu dalam kelompok yang lain. Dengan n_u dan n_v merupakan jumlah pengamatan dalam kelompok ke- u dan ke- v , ukuran jarak yang digunakan adalah

$$d_{w(uv)} = \frac{n_u}{n_u + n_v} d_{wu} + \frac{n_v}{n_u + n_v} d_{wv} \quad (2.10)$$

Menentukan jumlah kelompok optimum merupakan tahapan penting setelah proses pengelompokan. Tahapan ini disebut sebagai validasi pengelompokan (Halkidi, Batistakis, & Vizirgiannis, 2001). Indeks *R-squared* merupakan salah satu indeks yang dapat digunakan untuk menentukan jumlah kelompok optimum pada pengelompokkan hirarki (Sharma, 1996). Indeks tersebut melibatkan perhitungan keragaman data baik keragaman total, keragaman dalam kelompok, maupun keragaman antar

kelompok. Indeks validitas untuk menentukan jumlah kelompok optimum pada pengelompokan hirarki dapat dituliskan sebagai berikut

Sum of Square Total (SST)

$$SST = \sum_{l=1}^{m_{\text{numerik}}} \sum_{i=1}^n (x_{il} - \bar{x}_l)^2 \quad (2.11)$$

Sum of Square Within Group (SSW)

$$SSW = \sum_{c=1}^C \sum_{l=1}^{m_{\text{numerik}}} \sum_{i=1}^{n_c} (x_{ilc} - \bar{x}_{lc})^2 \quad (2.12)$$

Sum of Square Between Group (SSB)

$$SSB = SST - SSW \quad (2.13)$$

dengan,

m_{numerik} : Jumlah variabel numerik dalam pengamatan,

C : Jumlah kelompok yang dibentuk dalam pengamatan,

n : Total jumlah objek pengamatan,

n_c : Jumlah anggota pada kelompok ke- c untuk $c = 1, 2, \dots, C$

\bar{x}_l : Rata-rata keseluruhan objek pada variabel ke- l untuk

$l = 1, 2, \dots, m_{\text{numerik}}$,

\bar{x}_{lc} : Rata-rata variabel ke- l pada kelompok ke- c untuk

$c = 1, 2, \dots, C$.

R-squared kelompok baru adalah rasio dari *SSB* dan *SST*, artinya *R-squared* dapat didefinisikan sebagai ukuran perbedaan antar kelompok, dengan nilai berkisaran dari 0 sampai 1. Nilai *R-squared* = 0 menunjukkan bahwa tidak terdapat perbedaan antara kelompok, sedangkan *R-squared* = 1 menunjukkan adanya perbedaan yang signifikan antar kelompok yang terbentuk. Nilai *R-squared* dirumuskan pada persamaan 2.10

$$R - \text{squared} = \frac{SSB}{SST} = \frac{[SST - SSW]}{SST} \quad (2.10)$$

Penentuan jumlah kelompok yang terbentuk dapat dilihat berdasarkan nilai maksimum dari *Pseudo-F*. Rumus yang digunakan untuk menghitung nilai *pseudo F-statistics* adalah sebagai berikut.

$$PseudoF = \frac{\left(\frac{R^2}{c-1} \right)}{\left(\frac{1-R^2}{n-c} \right)} \quad (2.14)$$

Setelah mendapatkan jumlah cluster optimum dalam setiap tautan, selanjutnya dapat ditentukan metode tautan optimum untuk data numerik berdasarkan nilai *ICD Rate*. Nilai *ICD Rate* yang semakin kecil menunjukkan perbedaan keanggotaan setiap *cluster* yang artinya pengelompokan berdasarkan data yang sangat mirip dikelompokkan dalam satu *cluster* dan nilai SSE menunjukkan performansi *cluster* menghitung jumlah kuadrat *error* tiap metode. Berikut adalah rumus perhitungan *ICD Rate*.

$$ICDRate = 1 - \frac{SSB}{SST} = 1 - R^2 \quad (2.15)$$

2.4.2 Pengelompokan Data Kategorik

Pengelompokan data kategorik dilakukan dengan menggunakan ukuran kemiripan atau jarak untuk data kategorik kemudian dapat dilakukan pengelompokan dengan menggunakan metode hirarki maupun non hirarki. Namun demikian, metode hirarki dan non hirarki tradisional dinilai tidak tepat digunakan pada data kategorik sehingga dikembangkan metode ROCK (*RObust Clustering with linKs*) untuk mengelompokkan data kategorik tersebut (Guha, Rastogi, & Shim, 2000).

Metode ROCK menggunakan konsep *link* untuk mengukur kesamaan/kedekatan antara pengamatan. Jumlah *link* antar pengamatan bergantung pada nilai *threshold* (θ) yang telah ditentukan. Nilai θ merupakan parameter untuk menyatakan adanya *link* antar pengamatan. Jika jarak antar sepasang pengamatan lebih besar dari nilai θ , maka pengamatan tersebut memiliki link. Pengamatan yang memiliki tingkat hubungan (*link*) tinggi digabungkan ke dalam satu kelompok, sedang yang

mempunyai tingkat hubungan (*link*) yang kecil dipisahkan dari kelompok dimana data tersebut dikelompokkan.

Pengelompokan data kategorik dengan algoritma ROCK dilakukan dengan tiga langkah. Langkah pertama adalah menghitung similaritas. Ukuran kemiripan antara pasangan objek ke- i dan objek ke- j dihitung dengan persamaan 2.16

$$sim(X_i, X_j) = \frac{|X_i \cap X_j|}{|X_i \cup X_j|}, i \neq j \quad (2.16)$$

dengan,

$i = 1, 2, 3, \dots, n$ dan $j = 1, 2, 3, \dots, n$

X_i : Himpunan pengamatan ke- i dengan $X_i = \{x_{1i}, x_{2i}, \dots, x_{m_{kategorik\ i}}\}$,

X_j Himpunan pengamatan ke- j dengan

$X_j = \{x_{1j}, x_{2j}, \dots, x_{m_{kategorik\ j}}\}$,

$|X|$: Bilangan kardinal atau jumlah anggota dari himpunan X .

Langkah kedua adalah menentukan tetangga, pengamatan dinyatakan sebagai tetangga jika nilai $sim(X_i, X_j) \geq \theta$. Langkah terakhir adalah menghitung *link* antar objek pengamatan. Besarnya *link* dipengaruhi oleh nilai *threshold* (θ) yang digunakan untuk mengontrol seberapa dekat hubungan antar objek. Besarnya nilai θ yang diinputkan adalah $0 < \theta < 1$. Algoritma ROCK berhenti ketika jumlah dari kelompok yang diharapkan sudah terpenuhi atau tidak ada lagi *link* antara kelompok-kelompok (Dutta, Mahanta, & Arun, 2005).

Metode ROCK menggunakan informasi tentang *link* sebagai ukuran kemiripan antar objek. Jika terdapat objek pengamatan X_i , X_j dan X_k dimana X_i tetangga dari X_j , dan X_j tetangga dari X_k maka dikatakan X_i memiliki *link* dengan X_k walaupun X_i bukan tetangga dari X_k . Cara untuk menghitung *link* untuk semua kemungkinan pasangan dari n objek dapat menggunakan matriks **A**. Matriks **A** merupakan matriks

berukuran $n \times n$ yang bernilai 1 jika X_i dan X_j dinyatakan mirip (tetangga) dan bernilai 0 jika X_i dan X_j tidak mirip (bukan tetangga). Jumlah *link* antar pasangan X_i dan X_j diperoleh dari hasil kali antara baris ke X_i dan kolom ke X_j dari matriks **A**. Jika *link* antara X_i dan X_j semakin besar maka semakin besar kemungkinan X_i dan X_j berada dalam suatu kelompok yang sama (Dutta, Mahanta, & Arun, 2005).

Penggabungan kelompok dengan menggunakan algoritma ROCK didasarkan atas ukuran kebaikan (*goodness measure*) antar kelompok sebagaimana yang dituliskan dalam persamaan 2.17. *Goodness measure* adalah persamaan yang jumlah *link* dibagi dengan kemungkinan *link* yang terbentuk berdasarkan ukuran kelompoknya (Tyagi & Sharma, 2012).

$$g(C_i, C_j) = \frac{\text{link}[C_i, C_j]}{(n_i + n_j)^{1+2f(\theta)} - n_i^{1+2f(\theta)} - n_j^{1+2f(\theta)}} \quad (2.17)$$

dengan $\text{link}[C_i, C_j] = \sum_{X_i \in C_i, X_j \in C_j} \text{link}(X_i, X_j)$ yang menyatakan jumlah semua kemungkinan pasangan objek yang ada di dalam C_i dan C_j , serta n_i dan n_j masing-masing menyatakan jumlah anggota dalam kelompok ke- i dan j . Adapun $f(\theta)$ dinyatakan sebagai $\frac{1-\theta}{1+\theta}$.

2.5 Pengelompokan Data Campuran Numerik dan Kategorik

Analisis pengelompokan terhadap data campuran diawali dengan pembagian data menjadi dua subdata, yaitu data murni numerik dan data murni kategorik. Misalkan terdapat data dengan variabel berskala campuran sebanyak m , dengan m_{numerik} merupakan jumlah variabel yang murni berskala numerik dan

$m_{\text{kategorik}}$ merupakan jumlah variabel murni yang berskala kategorik, sehingga diperoleh $m = m_{\text{numerik}} + m_{\text{kategorik}}$. Selanjutnya dilakukan pengelompokan data sesuai dengan jenis data secara terpisah. Hasil pengelompokan tersebut kemudian digabungkan menggunakan metode pengelompokan ensemble ROCK sehingga diperoleh kelompok akhir (*final cluster*).

Pengelompokan ensemble merupakan metode yang menggabungkan beberapa algoritma yang berbeda untuk mendapatkan partisi umum dari data yang bertujuan untuk menyatukan hasil pengelompokan individu (Suguna & Selvi, 2012). Pengelompokan ensemble terdiri dari dua tahap algoritma. Tahap pertama adalah melakukan pengelompokan dengan beberapa algoritma dan menyimpan hasil pengelompokan tersebut. Tahap kedua menggunakan fungsi konsensus untuk menentukan *final cluster* dari kelompok-kelompok hasil tahap pertama.

Skema pengelompokan ensemble untuk data campuran dapat menggunakan algoritma CEBMDC (*Cluster Ensemble Based Mixed Data Clustering*). Langkah-langkah dalam analisis data campuran menggunakan metode pengelompokan ensemble yang disebut algoritma CEBMDC memiliki tahapan sebagai berikut (He, Xu, & Deng, 2005b).

- a. Membagi data menjadi dua subdata, yaitu murni numerik dan murni kategorik.
- b. Melakukan pengelompokan objek yang memiliki variabel numerik dengan algoritma pengelompokan data numerik, serta melakukan pengelompokan objek yang memiliki variabel kategorik dengan algoritma pengelompokan data kategorik.
- c. Menggabungkan (*combining*) hasil pengelompokan dari variabel numerik dan kategorik, yang disebut proses ensemble.

- d. Melakukan pengelompokan ensemble menggunakan algoritma pengelompokan data kategorik untuk mendapatkan kelompok akhir (*final cluster*).

2.6 Kinerja Hasil Pengelompokan

Pengukuran kinerja hasil pengelompokan merupakan langkah untuk mengetahui validitas suatu pengelompokan. Kinerja hasil pengelompokan untuk variabel dengan skala numerik dapat diketahui dari rasio nilai S_W dan S_B (Bunkers & James, 1996). Dengan menggunakan nilai rata-rata variabel, simpangan baku di dalam kelompok (S_W) dan simpangan baku antar kelompok (S_B) dapat dirumuskan dalam persamaan 2.18 dan 2.19

$$S_W = \frac{1}{C} \sum_{c=1}^C S_c \quad (2.18)$$

dengan S_c merupakan simpangan baku kelompok ke- c dan C adalah jumlah kelompok yang terbentuk.

$$S_B = \left[\frac{1}{C-1} \sum_{c=1}^C (\bar{x}_c - \bar{x})^2 \right]^{1/2} \quad (2.19)$$

dimana \bar{x}_c merupakan rata-rata kelompok ke- c dan \bar{x} adalah rata-rata keseluruhan kelompok. Kinerja suatu metode pengelompokan akan semakin baik jika semakin kecil nilai rasio antara S_W dan S_B . Hal ini berarti terdapat homogenitas maksimum dalam kelompok dan heterogenitas maksimum antara kelompok (Bunkers & James, 1996).

Adapun untuk mengukur kinerja pengelompokan untuk data yang kategorik adalah dengan menggunakan tabel kontingensi yang ekuivalen dengan melakukan ANOVA (*Analysis of Variance*). Jika terdapat sebanyak n pengamatan dengan n_k merupakan jumlah pengamatan dengan kategori ke- k dimana $k = 1, 2, 3, \dots, K$ dan $\sum_{k=1}^K n_k = n$. Selanjutnya n_{kc} merupakan jumlah pengamatan dengan kategori ke- k dan kelompok ke- c , dimana $c = 1, 2, 3, \dots, C$ dengan C adalah jumlah kelompok yang terbentuk, sehingga

$n_{.c} = \sum_{k=1}^K n_{kc}$ merupakan jumlah pengamatan pada kelompok ke- c

dan $n_{k.} = \sum_{c=1}^C n_{kc}$ merupakan jumlah pengamatan pada kategori ke- k .

Total jumlah pengamatan dapat dituliskan menjadi

$$n = \sum_{c=1}^C n_{.c} = \sum_{k=1}^K n_{k.} = \sum_{k=1}^K \sum_{c=1}^C n_{kc}.$$

Jumlah kuadrat total (SST) untuk variabel dengan data kategorik dapat dirumuskan seperti persamaan 2.20. Untuk total jumlah kuadrat dalam kelompok (SSW) dirumuskan dalam persamaan 2.21. Serta untuk jumlah kuadrat antar kelompok (SSB) dapat dirumuskan seperti pada persamaan 2.22 (Dewi, 2012).

$$SST = \frac{n}{2} - \frac{1}{2n} \sum_{k=1}^K n_k^2 \quad (2.20)$$

$$SSW = \sum_{c=1}^C \left(\frac{n_{.c}}{2} - \frac{1}{2n_{.c}} \sum_{k=1}^K n_{kc}^2 \right) = \frac{n}{2} - \frac{1}{2} \sum_{c=1}^C \frac{1}{n_{.c}} \sum_{k=1}^K n_{kc}^2 \quad (2.21)$$

$$SSB = \frac{1}{2} \left(\sum_{c=1}^C \frac{1}{n_{.c}} \sum_{k=1}^K n_{kc}^2 \right) - \frac{1}{2n} \sum_{k=1}^K n_k^2. \quad (2.22)$$

Mean of Squares Total (MST), *Mean of Squares Within (MSW)* dan *Mean of Squares Between (MSB)* dapat dirumuskan seperti pada persamaan 2.23 , 2.24 dan 2.25.

$$MST = \frac{SST}{(n-1)} \quad (2.23)$$

$$MSW = \frac{SSW}{(n-C)} \quad (2.24)$$

$$MSB = \frac{SSB}{C-1} \quad (2.25)$$

Simpangan baku dalam kelompok (S_w) dan simpangan baku antar kelompok (S_b) untuk data kategorik dapat dirumuskan pada persamaan 2.26 dan 2.27.

$$S_w = [MSW]^{1/2} \quad (2.26)$$

$$S_B = [MSB]^{1/2} \quad (2.27)$$

Selanjutnya, seperti halnya data numerik, kinerja suatu pengelompokan dengan data kategorik juga didasarkan pada perbandingan rasio antara simpangan baku dalam kelompok (S_w) dan simpangan baku antar kelompok (S_B). Dimana jika rasio perbandingan semakin kecil maka kinerja pengelompokan data kategorik semakin baik karena homogenitas maksimum dalam kelompok dan heterogenitas maksimum dalam kelompok.

2.7 Two Step Cluster

Two step cluster adalah alat eksplorasi yang dirancang untuk mengungkapkan sifat alami pengelompokan dalam kumpulan data yang seharusnya tidak terlihat. Two step cluster dapat menangani variabel numerik dan kategorik. Two step cluster mengasumsikan setiap variabel dalam cluster adalah independen. Selain itu variabel numerik diasumsikan mengikuti distribusi normal sedangkan variabel kategorik diasumsikan mengikuti distribusi multinomial. Two step cluster terdiri dari dua langkah, yaitu Pre-clustering dan Clustering step.

1. Langkah pre-clustering dimulai dengan memeriksa data awal untuk memutuskan apakah masuk ke dalam cluster yang telah terbentuk atau memulai cluster baru berdasarkan kriteria jarak. Metode ini dapat menggunakan dua tipe pengukuran jarak yaitu Euclidean dan Loglikelihood. Euclidean dapat digunakan untuk data kategorik sedangkan Loglikelihood dapat digunakan baik pada data numerik dan kategorik. Jarak Log likelihood dapat dirumuskan sebagai berikut.

$$d_{(i,j)} = \xi_i + \xi_j - \xi_{(i,j)} \quad (2.28)$$

Pre-clustering menggunakan konsep Clustering Feature (CF) dalam mengelompokan. Di CF terdapat node dan setiap node memiliki entri. Pada tahap ini memeriksa daun terdekat untuk masuk ke dalam node daun. Jika nilainya berdekatan dengan

threshold maka akan masuk ke dalam node daun, namun jika tidak maka akan ditentukan nilai baru pada node daun.

2. Pada langkah clustering, subcluster diperoleh dari proses pre-clustering yang digunakan sebagai input untuk kemudian dikelompokkan sesuai jumlah cluster yang diinginkan. Pada metode ini juga tidak dibutuhkan input parameter yang spesifik seperti jumlah cluster, karena metode ini secara otomatis menggunakan BIC dan kriteria informasi AIC. Sehingga estimasi awal mengenai jumlah cluster dihitung dengan mudah menggunakan indikator ini.

Two step cluster digunakan dalam penelitian ini sebagai metode pembandingan dalam mengelompokkan data campuran berskala numerik dan kategorik.

2.8 One Way ANOVA

ANOVA digunakan untuk membandingkan rata-rata dari beberapa populasi yang diwakili oleh beberapa kelompok sampel secara bersama, sehingga hipotesis matematikanya adalah

$$H_0 : \bar{\mu}_1 = \bar{\mu}_2 = \dots = \bar{\mu}_C = \bar{\mu}$$

$$H_1 : \text{minimal ada satu } \bar{\mu}_i \text{ yang tidak sama, dimana } i = 1, 2, \dots, C$$

Dengan statistik uji yaitu

$$F = \frac{MSB}{MSW} \quad (2.29)$$

Dimana MSB merupakan MSW diperoleh melalui persamaan

$$MSB = \frac{(SST - SSW)}{k - 1} \quad (2.30)$$

$$MSW = \frac{(SSW)}{N - k} \quad (2.31)$$

Adapun SST dan SSW diperoleh dari persamaan 2.11 dan 2.12. Daerah kritis dalam pengujian ini adalah Tolak H_0 jika

$$F > F_{k-1, N-k, \frac{\alpha}{2}}$$

2.9 One Way MANOVA

Setelah mendapatkan hasil pengelompokan yang optimum, berikutnya dilakukan uji perbedaan rata-rata dari kelompok yang diperoleh. Pada data multivariat pengujian rata-rata tersebut dilakukan dengan MANOVA (*Multivariate Analysis of Variance*). MANOVA merupakan perluasan dari teknik univariat *Analysis of Variance* (ANOVA) yang melibatkan lebih dari satu variabel (Johnson & Wicherin, 2007). Rumusan hipotesis pada MANOVA adalah sebagai berikut,

$$H_0 : \bar{\mu}_1 = \bar{\mu}_2 = \dots = \bar{\mu}_C = \bar{\mu}$$

$$H_1 : \text{minimal ada satu } \bar{\mu}_i \text{ yang tidak sama, dimana } i = 1, 2, \dots, C$$

Statistik uji yang digunakan dalam pengambilan keputusan dalam perbedaan antar kelompok adalah *wilk's lambda*. Nilai statistik uji *wilk's lambda* berkisar antara 0 sampai 1. Semakin rendah nilai statistik *wilk's lambda*, maka perbedaan antara kelompok semakin signifikan. Rumusan untuk statistik uji *wilk's lambda*, didefinisikan pada persamaan 2.32

$$\Lambda^* = \frac{|W|}{|B + W|} \quad (2.32)$$

Tabel 2.3 Tabel Transformasi Nilai Wilk's Lambda

Jumlah Variabel (m_{numerik})	Jumlah Kelompok	Distribusi Sampling
$m_{\text{numerik}} = 1$	$C \geq 2$	$\left(\frac{\sum n_c - C}{C - 1} \right) \left(\frac{1 - \Lambda^*}{\Lambda^*} \right) \sim F_{C-1, \sum n_c - C}$
$m_{\text{numerik}} = 2$	$C \geq 2$	$\left(\frac{\sum n_c - C - 1}{C - 1} \right) \left(\frac{1 - \sqrt{\Lambda^*}}{\Lambda^*} \right) \sim F_{2(C-1), 2(\sum n_c - C - 1)}$
$m_{\text{numerik}} \geq 1$	$C = 2$	$\left(\frac{\sum n_c - m_{\text{numerik}} - 1}{m_{\text{numerik}}} \right) \left(\frac{1 - \Lambda^*}{\Lambda^*} \right) \sim F_{m_{\text{numerik}}, \sum n_c - m_{\text{numerik}} - 1}$
$m_{\text{numerik}} \geq 1$	$C = 3$	$\left(\frac{\sum n_c - m_{\text{numerik}} - 2}{m_{\text{numerik}}} \right) \left(\frac{1 - \sqrt{\Lambda^*}}{\Lambda^*} \right) \sim F_{2m_{\text{numerik}}, 2(\sum n_c - m_{\text{numerik}} - 2)}$

Dimana $|W|$ merupakan silai SSW dengan rumusan seperti persamaan 2.12 dan $|B+W|$ merupakan matriks nilai SST dengan rumusan pada persamaan 2.11. Nilai statistik ujiwilk's lambda ditransformasi menjadi statistik F sehingga dapat dilakukan perbandingan dengan tabel F . Bentuk transformasi disajikan dalam Tabel 2.3

Dalam melakukan analisis dengan metode MANOVA, data yang dimiliki harus berdistribusi normal multivariat dan terdapat hubungan antar variabel (dependen). Terdapat beberapa teknik atau metode pemeriksaan asumsi distribusi normal multivariat, diantaranya yaitu metode proporsi dan metode plot Chi-Square. Sedangkan untuk menguji asumsi distribusi normal multivariat digunakan metode koefisien korelasi atau metode linieritas plot Chi-Square dan *mshapiro test* Pada penelitian ini akan lebih dijelaskan mengenai uji asumsi distribusi normal multivariat dengan menggunakan metode *mshapiro test*.

Metode *mshapiro test* ini berfungsi untuk menguji apakah data berdistribusi normal multivariat. Hipotesis, statistik uji, dan keputusan yang digunakan sebagai berikut.

H_0 : Data berdistribusi normal multivariat

H_1 : Data tidak berdistribusi normal multivariat

Statistik uji untuk pengujian asumsi berdistribusi normal multivariat ditunjukkan dalam persamaan 2.33

$$W = \frac{\left(\sum_{i=1}^n a_i x_{(i)} \right)^2}{\sum_{i=1}^n (x_i - \bar{x})^2} \quad (2.33)$$

Dimana $(a_1, a_2, \dots, a_n) = \frac{m^T V^{-1}}{(m^T V^{-1} V^{-1} m)^{1/2}}$. Diambil keputusan tolak

H_0 dengan ditetapkannya taraf signifikan sebesar α maka H_0 ditolak jika W kurang dari a .

Untuk menguji hubungan antar variabel digunakan Uji Bartlett. Variabel dikatakan saling bebas (independen) jika matriks

korelasi antar variabel membentuk matriks identitas. Untuk menguji kebebasan antar variabel ini dapat dilakukan dengan menggunakan uji *Bartlett Sphericity* dengan uji hipotesis sebagai berikut.

$$H_0 : \rho = I$$

$$H_1 : \rho \neq I$$

Dengan statistik uji didapatkan dari perhitungan berikut.

$$\chi^2_{hitung} = - \left\{ n - 1 - \frac{2p+5}{6} \right\} \ln |\mathbf{R}| \quad (2.34)$$

Daerah kritis dari pengujian ini yaitu H_0 ditolak jika $\chi^2_{hitung} > \chi^2_{\alpha, \frac{1}{2}p(p-1)}$

2.10 Uji Mann Whitney

Uji Mann-Whitney digunakan untuk melihat kesamaan parameter-parameter lokasi populasi pada sampel yang memiliki ukuran tidak sama. Asumsi yang diperlukan dalam pengujian Mann-Whitney adalah data merupakan sampel acak hasil pengamatan dari 2 populasi, kedua sampel tidak saling mempengaruhi, dan skala pengukuran yang dipakai sekurang-kurangnya adalah ordinal (Daniel, 1989).

Hipotesis dalam pengujian Mann-Whitney adalah

H_0 : Kedua populasi identik

H_1 : Kedua populasi berbeda dalam hal lokasi

Dengan statistik uji didapatkan dari perhitungan berikut

$$T = S - \frac{n_1(n_1 + 1)}{2} \quad (2.35)$$

Dimana S adalah jumlah peringkat hasil-hasil pengamatan yang merupakan sampel dari populasi 1. Dengan daerah kritis yaitu tolak H_0 jika T kurang dari $W_{\alpha/2}$.

2.11 Desa Tertinggal

Desa adalah desa dan desa adat atau yang disebut dengan nama lain, selanjutnya disebut desa adalah kesatuan masyarakat

hukum yang memiliki batas wilayah yang berwenang untuk mengatur dan mengurus urusan pemerintahan, kepentingan masyarakat setempat berdasarkan prakarsa masyarakat, hak asal usul, dan/atau hak tradisional yang diakui dan dihormati dalam sistem pemerintahan Negara Kesatuan Republik Indonesia (Undang-undang no 6 tahun 2014 tentang Desa).

Pada tahun 2015, Badan Perencanaan Nasional dan Badan Pusat Statistik menerbitkan Indeks Pembangunan Desa (IPD) 2014 yang memuat klasifikasi desa. IPD merupakan indeks komposit tertimbang dari 42 indikator yang secara substansi dan bersama-sama menggambarkan tingkat pembangunan di desa. Setiap indikator harus memiliki kontribusi terhadap IPD. Besarnya kontribusi setiap indikator menggambarkan besarnya pengaruh indikator tersebut terhadap dimensi dan IPD. Besarnya kontribusi setiap indikator ditetapkan berdasarkan sebaran data menggunakan teknik statistik yaitu PCA. Penimbang setiap indikator penyusun IPD dihitung berdasarkan *loading factor* yang dihasilkan dari model PCA yang paling optimum. Besarnya bobot tiap indikator dalam faktor serta besarnya kontribusi tiap indikator dituliskan berturut-turut dalam persamaan 2.36 dan 2.37

$$B = \frac{LF}{RLF} \times RSSL \quad (2.36)$$

$$b = \frac{B}{JB} \quad (2.37)$$

Dimana

B = nilai bobot

LF = nilai *loading factor*

RLF = rata-rata *loading factor* dalam 1 faktor

$RSSL$ = nilai *rotation sum of square loading* (% variance)

b = nilai kontribusi

JB = jumlah semua bobot

Selanjutnya nilai IPD diperoleh dari penjumlahan secara tertimbang terhadap setiap indikator penyusun IPD. Nilai yang dijumlahkan merupakan skor setiap indikator yang sudah

ditimbang atau dikalikan dengan penimbang masing-masing indikator. Secara matematis dituliskan sebagai berikut

$$IPD = (b_1V_1 + b_2V_2 + \dots + b_{42}V_{42}) \times 20 \quad (2.38)$$

Dimana

IPD = nilai Indeks Pembangunan Desa (bernilai 0-100)

V_1 = skor indikator ke-1

V_2 = skor indikator ke-2

\vdots

V_{42} = skor indikator ke-42

b_1 = pembobot indikator ke-1

b_2 = pembobot indikator ke-2

\vdots

b_{42} = pembobot indikator ke-42

Berdasarkan IPD tahun 2014, desa dibagi menjadi tiga klasifikasi yaitu Desa Mandiri, Desa Berkembang, dan Desa Tertinggal. Desa Mandiri adalah desa yang telah terpenuhi SPM desa mencakup beberapa aspek yaitu kebutuhan sosial dasar, infrastruktur dasar, sarana dasar, pelayanan umum, dan penyelenggaraan pemerintahan desa, serta kelembagaan desa yang keberlanjutan. Desa Berkembang adalah desa yang sudah terpenuhi SPM desa pada semua aspek tetapi pengelolaannya belum menunjukkan keberlanjutan. Sedangkan, Desa Tertinggal adalah desa yang belum terpenuhi SPM desa pada aspek kebutuhan sosial dasar, infrastruktur dasar, sarana dasar, pelayanan umum, dan penyelenggaraan pemerintahan (Bappenas dan BPS, 2015).

IPD terdiri dari lima dimensi yang disesuaikan dengan ketersediaan data/variabel dalam data Potensi Desa 2014 yaitu sebagai berikut

1. Pelayanan Dasar mewakili aspek pelayanan dasar untuk mewujudkan bagian dari kebutuhan dasar, khusus untuk pendidikan dan kesehatan. Variabel yang termasuk sebagai komponen penyusunnya meliputi ketersediaan dan akses terhadap fasilitas pendidikan seperti TK, SD, SMP, dan SMA; serta ketersediaan dan akses terhadap fasilitas kesehatan seperti rumah sakit, rumah sakit bersalin,

puskesmas/pustu, tempat praktek dokter, poliklinik/balai pengobatan, tempat praktek bidan, poskesdes, polindes, dan apotek.

2. Kondisi Infrastruktur mewakili Kebutuhan Dasar; Sarana; Prasarana; Pengembangan Ekonomi Lokal; dan Pemanfaatan Sumberdaya Alam secara Berkelanjutan dengan memisahkan aspek aksesibilitas/transportasi. Variabel-variabel penyusunnya mencakup ketersediaan infrastruktur ekonomi seperti: kelompok pertokoan, minimarket, maupun toko kelontong, pasar, restoran, rumah makan, maupun warung/kedai makanan, akomodasi hotel atau penginapan, serta bank; ketersediaan infrastruktur energi seperti: listrik, penerangan jalan, dan bahan bakar untuk memasak; ketersediaan infrastruktur air bersih dan sanitasi seperti: sumber air minum, sumber air mandi/cuci, dan fasilitas buang air besar; serta ketersediaan dan kualitas infrastruktur komunikasi dan informasi seperti: komunikasi menggunakan telepon seluler, internet, dan pengiriman pos/barang.
3. Aksesibilitas/Transformasi dipisahkan sebagai dimensi tersendiri dalam indikator pembangunan desa dengan pertimbangan sarana dan prasarana transportasi memiliki kekhususan dan prioritas pembangunan desa sebagai penghubung kegiatan sosial ekonomi dalam desa. Variabel-variabel penyusunnya meliputi ketersediaan dan akses terhadap sarana transportasi seperti: lalu lintas dan kualitas jalan, aksesibilitas jalan, ketersediaan dan operasional angkutan umum; dan aksesibilitas transportasi seperti: waktu tempuh per kilometer transportasi ke kantor camat, biaya per kilometer transportasi ke kantor camat, waktu tempuh per kilometer transportasi ke kantor bupati/walikota, dan biaya per kilometer transportasi ke kantor bupati/walikota. Indikator jarak antara desa dengan pusat pemerintahan ini merujuk dari variabel yang tercantum dalam Potensi Desa 2014, dengan asumsi bahwa pada

umumnya pusat kegiatan sosial-ekonomi suatu kawasan berada di sekitar/dekat pusat-pusat pemerintahan.

4. Pelayanan Umum merupakan upaya pemenuhan kebutuhan pelayanan atas barang, jasa, dan/atau pelayanan administratif dengan tujuan memperkuat demokrasi, kohesi sosial, perlindungan lingkungan, dan sebagainya. Pelayanan dalam dimensi ini mewakili aspek lingkungan dan aspek pemberdayaan masyarakat serta mengacu pada ketersediaan data Potensi Desa 2014. Aspek lingkungan dalam hal ini terkait dengan kesehatan lingkungan masyarakat, sedangkan aspek pemberdayaan masyarakat diwakili dengan keberadaan kelompok kegiatan masyarakat. Oleh karena itu, variabel-variabel penyusun dimensi ini mencakup penanganan kesehatan masyarakat seperti: penanganan kejadian luar biasa (KLB), dan penanganan gizi buruk; serta ketersediaan fasilitas olah raga seperti: ketersediaan lapangan olah raga, dan kelompok kegiatan olah raga.
5. Penyelenggaraan Pemerintahan mewakili indikasi kinerja pemerintahan desa merupakan bentuk pelayanan administratif yang diselenggarakan penyelenggara pelayanan bagi warga yang dalam hal ini adalah Pemerintah. Oleh karena itu variabel ini perlu diukur dan berdiri sendiri sebagai sebuah indikator pembangunan desa, karena sifatnya sebagai perangkat terlaksananya tujuan pembangunan desa tersebut. Variabel-variabel penyusunnya meliputi kemandirian seperti: kelengkapan pemerintahan desa, otonomi desa, dan asset/kekayaan desa; serta kualitas sumber daya manusia seperti: kualitas SDM kepala desa dan sekretaris desa.

IPD disusun untuk menunjukkan tingkat perkembangan pembangunan di suatu desa. Nilai indeks mempunyai rentang 0 s/d 100. Untuk memudahkan interpretasi, maka dilakukan pengelompokan desa menjadi 3 kategori yaitu

1. Desa Mandiri adalah desa yang telah terpenuhi SPM desa mencakup beberapa aspek yaitu kebutuhan sosial dasar,

infrastruktur dasar, sarana dasar, pelayanan umum, dan penyelenggaraan pemerintahan desa, serta kelembagaan desa yang keberlanjutan. Secara teknis desa mandiri adalah desa yang telah memiliki nilai IPD lebih dari 75.

2. Desa Berkembang adalah desa yang sudah terpenuhi SPM desa pada semua aspek tetapi pengelolaannya belum menunjukkan keberlanjutan. Secara teknis memiliki nilai IPD yang lebih dari 50 namun kurang dari 75.
3. Desa Tertinggal adalah adalah desa yang belum terpenuhi SPM desa pada aspek kebutuhan sosial dasar, infrastruktur dasar, sarana dasar, pelayanan umum, dan penyelenggaraan pemerintahan. Secara teknis memiliki nilai IPD dibawah 50.

2.12 Gambaran Umum Kabupaten Bondowoso

Kabupaten Bondowoso, adalah kabupaten di Provinsi Jawa Timur, Indonesia. Ibu kotanya adalah Bondowoso. Ibu kota kabupaten Bondowoso berada di persimpangan jalur dari Besuki dan Situbondo menuju Jember. Kabupaten Bondowoso merupakan salah satu kabupaten yang tidak memiliki wilayah laut (terkurung daratan) dan terletak di wilayah Tapal Kuda, Jawa Timur. Kabupaten Bondowoso memiliki luas wilayah 1.560,10 km² yang secara geografis berada pada koordinat antara 113°48'10" - 113°48'26" BT dan 7°50'10" - 7°56'41" LS. Batas wilayah Kabupaten Bondowoso sebelah utara dengan Kabupaten Situbondo, sebelah timur dengan Kabupaten Situbondo dan Kabupaten banyuwangi, sebelah selatan dengan Kabupaten Jember dan sebelah barat dengan Kabupaten Situbondo dan Kabupaten Probolinggo. Kabupaten Bondowoso dapat dibagi menjadi tiga wilayah: Wilayah barat merupakan pegunungan (bagian dari Pegunungan Iyang), bagian tengah berupa dataran tinggi dan bergelombang, sedang bagian timur berupa pegunungan (bagian dari Dataran Tinggi Ijen). Bondowoso merupakan satu-satunya kabupaten di daerah Tapal Kuda yang tidak memiliki garis pantai. Letak Kabupaten Bondowoso tidak berada pada daerah yang strategis. Meskipun berada di tengah, namun Kabupaten Bondowoso tidak dilalui jalan negara yang menghubungkan antar

provinsi. Bondowoso juga tidak memiliki lautan. Ini yang menyebabkan Bondowoso sulit berkembang dibandingkan dengan kabupaten lainnya di Jawa Timur.



Gambar 2.1 Peta Kabupaten Bondowoso

(Sumber : Bappeda Kabupaten Bondowoso)

Wilayah administrasi Kabupaten Bondowoso terdiri dari 23 kecamatan yaitu Kecamatan Binakal, Bondowoso, Botolinggo, Cermee, Curahdami, Grujugan, Jembersari Darus Sholah, Klabang, Maesan, Pakem, Prajekan, Pujer, Sempol, Sukosari, Sumberwringin, Taman Krocok, Tamanan, Tapen, Tegalampel, Tenggarang, Tlogosari, Wonosari dan Wringin. Dari 23 kecamatan tersebut terdapat 219 desa.

(Halaman ini sengaja dikosongkan)

BAB III

METODOLOGI PENELITIAN

3.1 Sumber Data

Data yang digunakan dalam penelitian ini adalah data sekunder yang diperoleh dari Badan Pusat Statistik. Data tersebut merupakan data hasil survey Potensi Desa (PODES) Provinsi Jawa Timur tahun 2014 yang disesuaikan dengan indikator-indikator indeks pembangunan desa. Survey dilakukan pada Bulan April 2014. Unit sampelnya adalah desa yang memenuhi kriteria yaitu memiliki wilayah dengan batas yang jelas, memiliki penduduk yang menetap di wilayah desa dan memiliki pemerintahan desa. Survey Potensi Desa terhadap Kabupaten Bondowoso dilakukan terhadap 219 desa/kelurahan.

3.2 Variabel Penelitian

Variabel yang digunakan dalam penelitian ini terdiri dari variabel yang berskala numerik dan kategorik. Terdapat total 21 variabel yang digunakan dalam penelitian ini yang terdiri dari 11 variabel berskala numerik dan 10 variabel berskala kategorik. Penentuan variabel penelitian didasarkan pada indikator-indikator dalam Indeks Pembangunan Desa (IPD) yang disusun oleh Bappenas dalam menetapkan daerah tertinggal di Indonesia. Karena daerah tertinggal yang ditetapkan oleh Bappenas masih dalam lingkup Kabupaten, maka variabel penelitian yang dipilih disesuaikan dengan indikator dalam IPD yang mengukur tingkat perkembangan hanya dalam lingkup Desa. Daftar variabel penelitian disajikan dalam Tabel 3.1.

Tabel 3.1 Variabel Penelitian

Data Numerik		
No	Keterangan	Variabel
1	Jumlah Pos Kesehatan Desa/Poskesdes (unit)	X ₁
2	Jumlah Tempat Praktik Bidan (unit)	X ₂
3	Jumlah Tempat Praktik Dokter (unit)	X ₃
4	Jumlah Pasar dengan Bangunan Permanen (unit)	X ₄

Tabel 3.1 Variabel Penelitian (lanjutan)

No	Keterangan	Variabel
5	Jumlah Toko/Warung Kelontng (Unit)	X ₅
6	Persentase Keluarga Pengguna Listrik PLN (%)	X ₆
7	Nilai Pendapatan Asli Desa (Jutaan Rupiah)	X ₇
8	Nilai Alokasi Dana Desa (Jutaan Rupiah)	X ₈
9	Jumlah Sarana Pendidikan Sekolah Dasar/SD (unit)	X ₉
10	Jumlah Sarana Pendidikan Sekolah Menengah Pertama/SMP (unit)	X ₁₀
11	Jarak menuju Sarana Pendidikan Sekolah Menengah Atas/SMA terdekat (km)	X ₁₁
Data Kategorik		
No	Keterangan	Variabel
12	Kondisi Penerangan di Jalan Utama	X ₁₂
13	Bahan Bakar untuk Memasak	X ₁₃
14	Sumber Air untuk Minum	X ₁₄
15	Sumber Air untuk Mandi/Mencuci	X ₁₅
16	Fasilitas Buang Air Besar	X ₁₆
17	Kualitas Fasilitas Komunikasi Seluler	X ₁₇
18	Kelengkapan Pemerintahan Desa	X ₁₈
19	Aset/Kekayaan Desa	X ₁₉
20	Tingkat Pendidikan Kepala Desa	X ₂₀
21	Tingkat Pendidikan Sekretaris Desa	X ₂₁

Definisi operasional dari variabel penelitian sebagai berikut :

1. Pos Kesehatan Desa (Poskesdes) atau lebih sering dikenal sebagai PKD di beberapa wilayah merupakan sarana kesehatan/bangunan yang dibentuk di desa/kelurahan dalam rangka mendekatkan/menyediakan pelayanan kesehatan dasar bagi masyarakat desa/kelurahan.
2. Tempat praktek bidan adalah sarana kesehatan dan/atau bangunan yang digunakan untuk tempat praktek bidan yang biasanya memberikan pelayanan ibu hamil dan bayi.
3. Tempat praktek dokter adalah sarana kesehatan dan/atau bangunan yang digunakan untuk tempat praktek dokter yang biasanya memberikan pelayanan berobat jalan, termasuk

praktek dokter yang mempunyai fasilitas rawat inap dan apotek.

4. Pasar adalah tempat pertemuan antara penjual dan pembeli barang dan jasa. Pasar dengan bangunan permanen adalah pasar yang dengan kriteria memiliki atap, lantai dan dinding.
5. Toko/warung kelontong adalah bangunan yang berfungsi sebagai tempat usaha di bangunan tetap untuk menjual barang keperluan sehari-hari secara eceran, tidak mempunyai sistem pelayanan mandiri dan dikelola oleh satu penjual.
6. Persentase keluarga pengguna listrik PLN adalah persentase dari keluarga pengguna/pelanggan listrik yang disalurkan oleh PLN.
7. Pendapatan Asli Desa (PAD) adalah semua penerimaan uang melalui rekening desa yang merupakan hak desa dalam 1 (satu) tahun anggaran yang tidak perlu dibayar kembali oleh desa, yang terdiri dari hasil usaha, aset, swadaya partisipasi dan gotong royong serta lain-lain pendapatan asli desa.
8. Alokasi Dana Desa (ADD) adalah anggaran keuangan yang diberikan pemerintah kepada desa, yang mana sumbernya berasal dari Bagi Hasil Pajak Daerah serta dari Dana Perimbangan Keuangan Pusat Dan Daerah yang diterima oleh kabupaten.
9. Sarana pendidikan dasar meliputi Sekolah Dasar (SD) atau Madrasah Ibtidaiyah (MI) baik negeri maupun swasta yang terdapat dalam wilayah desa/kelurahan.
10. Sarana pendidikan menengah pertama meliputi Sekolah Menengah Pertama (SMP) atau Madrasah Tsanawiyah (MTS) baik negeri maupun swasta yang terdapat dalam wilayah desa/kelurahan.
11. Sarana pendidikan menengah atas meliputi Sekolah Menengah Atas (SMA), Sekolah Menengah Kejuruan (SMK) dan Madrasah Aliyah (MA) baik negeri atau swasta.

12. Kondisi penerangan di jalan utama dibedakan menjadi tidak ada penerangan, penerangan non listrik, dan penerangan listrik.

Kategori dari Bappenas :

- 1) **Kode 0 (nol)** merupakan desa yang tidak ada penerangan di jalan utama desa.
 - 2) **Kode 2 (dua)** merupakan desa yang ada penerangan di jalan utama desa dengan jenis penerangannya berupa non listrik.
 - 3) **Kode 3 (tiga)** merupakan desa yang ada penerangan di jalan utama desa dengan jenis penerangannya berupa listrik yang diusahakan oleh non pemerintah.
 - 4) **Kode 5 (lima)** merupakan desa yang ada penerangan di jalan utama desa dengan jenis penerangannya berupa listrik yang diusahakan oleh pemerintah.
13. Bahan bakar untuk memasak mengacu pada jenis bahan bakar untuk memasak yang digunakan oleh mayoritas keluarga di desa.

Kategori dari Bappenas :

- 1) **Kode 0 (nol)** merupakan desa yang sebagian besar keluarganya menggunakan batubara, arang, dan lain-lain untuk memasak.
- 2) **Kode 1 (satu)** merupakan desa yang sebagian besar keluarganya menggunakan kayu bakar untuk memasak.
- 3) **Kode 2 (dua)** merupakan desa yang sebagian besar keluarganya menggunakan minyak tanah untuk memasak, tetapi tidak ada agen/penjual minyak tanah.
- 4) **Kode 3 (tiga)** merupakan desa yang sebagian besar keluarganya menggunakan minyak tanah untuk memasak, dan ada agen/penjual minyak tanah.
- 5) **Kode 4 (empat)** merupakan desa yang sebagian besar keluarganya menggunakan LPG atau gas kota untuk memasak, tetapi tidak ada pangkalan atau agen atau penjual LPG.

- 6) **Kode 5 (lima)** merupakan desa yang sebagian besar keluarganya menggunakan LPG atau gas kota, dan ada pangkalan atau agen atau penjual LPG.
14. Sumber air minum mengacu kepada jenis sumber air minum yang digunakan oleh mayoritas keluarga di desa.
Kategori dari Bappenas :
- 1) **Kode 0 (nol)** merupakan desa yang sumber air untuk minum sebagian besar keluarga berasal dari air hujan atau lainnya.
 - 2) **Kode 1 (satu)** merupakan desa yang sumber air untuk minum sebagian besar keluarga berasal dari sungai/danau/kolam.
 - 3) **Kode 2 (dua)** merupakan desa yang sumber air untuk minum sebagian besar keluarga berasal dari mata air.
 - 4) **Kode 3 (tiga)** merupakan desa yang sumber air untuk minum sebagian besar keluarga berasal dari sumur.
 - 5) **Kode 4 (empat)** merupakan desa yang sumber air untuk minum sebagian besar keluarga berasal dari ledeng tanpa meteran, sumur bor, atau pompa.
 - 6) **Kode 5 (lima)** merupakan desa yang sumber air untuk minum sebagian besar keluarga berasal dari air kemasan, atau ledeng dengan meteran.
15. Sumber air untuk mandi/cuci mengacu kepada jenis sumber air yang digunakan oleh mayoritas keluarga di desa untuk mandi/cuci.
Kategori dari Bappenas :
- 1) **Kode 0 (nol)** merupakan desa yang sumber air untuk mandi/mencuci sebagian besar keluarga berasal dari air hujan atau lainnya.
 - 2) **Kode 1 (satu)** merupakan desa yang sumber air untuk mandi/mencuci sebagian besar keluarga berasal dari sungai/danau/kolam.
 - 3) **Kode 2 (dua)** merupakan desa yang sumber air untuk mandi/mencuci sebagian besar keluarga berasal dari mata air.

- 4) **Kode 3 (tiga)** merupakan desa yang sumber air untuk mandi/mencuci sebagian besar keluarga berasal dari sumur.
 - 5) **Kode 4 (empat)** merupakan desa yang sumber air untuk mandi/mencuci sebagian besar keluarga berasal dari ledeng tanpa meteran, sumur bor, atau pompa.
 - 6) **Kode 5 (lima)** merupakan desa yang sumber air untuk mandi/mencuci sebagian besar keluarga berasal dari air kemasan, atau ledeng dengan meteran.
16. Fasilitas buang air besar mengacu pada jenis fasilitas yang digunakan oleh mayoritas keluarga di desa
- Kategori dari Bappenas :
- 1) **Kode 0 (nol)** merupakan desa yang fasilitas buang air besar sebagian besar keluarga adalah bukan jamban.
 - 2) **Kode 2 (dua)** merupakan desa yang fasilitas buang air besar sebagian besar keluarga adalah jamban umum.
 - 3) **Kode 3 (tiga)** merupakan desa yang fasilitas buang air besar sebagian besar keluarga adalah jamban bersama.
 - 4) **Kode 5 (lima)** merupakan desa yang fasilitas buang air besar sebagian besar keluarga adalah jamban sendiri.
17. Kualitas fasilitas komunikasi seluler ditandai dengan adanya sinyal telepon seluler. Sinyal telepon seluler adalah besaran elektromagnetik yang berubah dalam ruang dan waktu dengan membawa informasi yang memberikan konfirmasi bahwa layanan telepon seluler sudah tersedia. Rincian ini terdiri dari tidak ada sinyal, sinyal lemah, dan sinyal kuat.
- Kategori dari Bappenas :
- 1) **Kode 0 (nol)** merupakan desa yang tidak ada *Base Transceiver Station* (BTS) dan tidak ada sinyal telepon seluler/*handphone*.
 - 2) **Kode 1 (satu)** merupakan desa yang ada *Base Transceiver Station* (BTS), tetapi tidak ada sinyal telepon seluler/*handphone*.

- 3) **Kode 2 (dua)** merupakan desa yang tidak ada *Base Transceiver Station* (BTS), tetapi ada sinyal telepon seluler/*handphone* yang lemah.
 - 4) **Kode 3 (tiga)** merupakan desa yang ada *Base Transceiver Station* (BTS), tetapi ada sinyal telepon seluler/*handphone* yang lemah.
 - 5) **Kode 4 (empat)** merupakan desa yang tidak ada *Base Transceiver Station* (BTS), tetapi ada sinyal telepon seluler/*handphone* yang kuat.
 - 6) **Kode 5 (lima)** merupakan desa yang ada *Base Transceiver Station* (BTS),tetapi ada sinyal telepon seluler/*handphone* yang kuat.
18. Kelengkapan pemerintahan desa mengacu pada keberadaan BPD dan kantor kepala desa.
- Kategori dari Bappenas
- 1) **Kode 0 (nol)** merupakan desa yang tidak ada BPD dan tidak ada kantor kepala desa. Tanpa mempertimbangkan ketersediaan batas wilayah desa dalam bentuk peta yang ditetapkan oleh bupati atau walikota.
 - 2) **Kode 1 (satu)** merupakan desa yang ada BPD tetapi tidak ada kantor kepala desa. Tanpa mempertimbangkan ketersediaan batas wilayah desa dalam bentuk peta yang ditetapkan oleh bupati atau walikota.
 - 3) **Kode 2 (dua)** merupakan desa yang tidak ada BPD tetapi ada kantor kepala desa di luar wilayah desa. Tanpa mempertimbangkan ketersediaan batas wilayah desa dalam bentuk peta yang ditetapkan oleh bupati/walikota.
 - 4) **Kode 3 (tiga)** merupakan desa yang ada BPD dan ada kantor kepala desa di luar wilayah desa. Tanpa mempertimbangkan ketersediaan batas wilayah desa dalam bentuk peta yang ditetapkan oleh bupati/walikota.
 - 5) **Kode 4 (empat)** merupakan desa yang tidak ada BPD tetapi ada kantor kepala desa di dalam wilayah desa. Tanpa mempertimbangkan ketersediaan batas wilayah

desa dalam bentuk peta yang ditetapkan oleh bupati/walikota.

- 6) **Kode 5 (lima)** merupakan desa yang ada BPD dan ada kantor kepala desa di dalam wilayah desa. Tanpa mempertimbangkan ketersediaan batas wilayah desa dalam bentuk peta yang ditetapkan oleh bupati/walikota.
19. Aset desa adalah barang milik desa yang berasal dari kekayaan asli desa, dibeli atau diperoleh atas beban Anggaran Pendapatan dan Belanja Desa atau perolehan hak lainnya yang sah. Aset desa terdiri dari Pendapatan Asli Desa (PAD), tanah kas desa/ulayat, bangunan desa (kantor kepala desa, balai desa, dan lain-lain), pasar desa (pasa hewan, pelelangan ikan, pelelangan hasil pertanian, dan lain-lain), maupun aset desa lainnya.

Kategori dari Bappenas

- 1) **Kode 0 (nol)** merupakan desa yang sama sekali tidak mempunyai kelima jenis aset/pendapatan desa
- 2) **Kode 1 (satu)** merupakan desa yang mempunyai 1 (satu) dari kelima jenis aset/pendapatan desa
- 3) **Kode 2 (dua)** merupakan desa yang mempunyai 2 (dua) dari kelima jenis aset/pendapatan
- 4) **Kode 3 (tiga)** merupakan desa yang mempunyai 3 (tiga) dari kelima jenis aset/pendapatan desa
- 5) **Kode 4 (empat)** merupakan desa yang mempunyai 4 (empat) dari kelima jenis aset/pendapatan desa
- 6) **Kode 5 (lima)** merupakan desa yang mempunyai 5 (lima) jenis aset/pendapatan desa.
20. Kualitas SDM Kepala Desa mengacu pada pendidikan tertinggi yang ditamatkan.

Kategori dari Bappenas:

- 1) **Kode 0 (nol)** merupakan desa yang tidak mempunyai kepala desa.
- 2) **Kode 1 (satu)** merupakan desa yang mempunyai kepala desa dengan pendidikan tertinggi tidak tamat SD sederajat atau tidak pernah sekolah.

- 3) **Kode 2 (dua)** merupakan desa yang mempunyai kepala desa dengan pendidikan tertinggi yang ditamatkannya adalah SD sederajat.
 - 4) **Kode 3 (tiga)** merupakan desa yang mempunyai kepala desa dengan pendidikan tertinggi yang ditamatkannya adalah SMP sederajat.
 - 5) **Kode 4 (empat)** merupakan desa yang mempunyai kepala desa dengan pendidikan tertinggi yang ditamatkannya adalah SMU sederajat.
 - 6) **Kode 5 (lima)** merupakan desa yang mempunyai kepala desa dengan pendidikan tertinggi yang ditamatkannya adalah akademi/DIII hingga S3.
21. Kualitas SDM Sekretaris Desa mengacu pada pendidikan tertinggi yang ditamatkan.
- Kategori dari Bappenas :
- 1) **Kode 0 (nol)** merupakan desa yang tidak mempunyai sekretaris desa.
 - 2) **Kode 1 (satu)** merupakan desa yang mempunyai sekretaris desa dengan pendidikan tertinggi tidak tamat SD sederajat atau tidak pernah sekolah.
 - 3) **Kode 2 (dua)** merupakan desa yang mempunyai sekretaris desa dengan pendidikan tertinggi yang ditamatkannya adalah SD sederajat.
 - 4) **Kode 3 (tiga)** merupakan desa yang mempunyai sekretaris desa dengan pendidikan tertinggi yang ditamatkannya adalah SMP sederajat.
 - 5) **Kode 4 (empat)** merupakan desa yang mempunyai sekretaris desa dengan pendidikan tertinggi yang ditamatkannya adalah SMU sederajat.
 - 6) **Kode 5 (lima)** merupakan desa yang mempunyai sekretaris desa dengan pendidikan tertinggi yang ditamatkannya adalah akademi/DIII hingga S3.

3.3 Langkah Penelitian

Dalam subbab ini terdiri dari langkah analisis untuk menyelesaikan tujuan pertama yaitu menentukan karakteristik desa

di Kabupaten Bondowoso berdasarkan indikator pembangunan desa, tujuan kedua dan ketiga yaitu melakukan pengelompokan desa di Kabupaten Bondowoso berdasarkan indikator pembangunan desa yang terdiri dari data berskala campuran dengan menggunakan metode ensembel ROCK dan *two step cluster*. Serta tujuan keempat yaitu mendapatkan hasil pengelompokan terbaik diantara metode ensembel ROCK dan *two step cluster*. Sebelumnya, dilakukan pembagian data dari 21 variabel berskala campuran, dipisahkan menjadi 11 variabel berskala numerik (X_1 - X_{11}) dan 10 variabel berskala kategorik (X_{12} - X_{21}). Selanjutnya dilakukan analisis untuk menyelesaikan tujuan.

Untuk menyelesaikan tujuan yang pertama, langkah analisis yang dilakukan sebagai berikut

1. Melakukan deskripsi data untuk variabel berskala numerik dengan menentukan nilai maksimum dan nilai minimum serta *mean* dari setiap variabel tersebut.
2. Melakukan deskripsi data untuk variabel berskala kategorik dengan menentukan modus dari setiap variabel tersebut.

Selanjutnya, untuk menyelesaikan tujuan yang kedua dan ketiga, langkah analisis yang dilakukan sebagai berikut.

1. Melakukan analisis faktor terhadap data variabel numerik untuk menghilangkan multikolineritas antar variabel numerik.
2. Melakukan pengelompokan untuk variabel yang berskala numerik dengan menggunakan metode hirarki agglomerative. Jarak yang digunakan adalah jarak Euclidean dan pautan yang digunakan adalah *single linkage*, *complete linkage* dan *average linkage*.
3. Melakukan perhitunga *Pseudo-F* dan *ICD Rate* untuk menentukan pautan yang digunakan dengan jumlah kelompok optimum untuk metode hirarki agglomerative yang selanjutnya dinyatakan sebagai *output* cluster numerik.
4. Melakukan pengelompokan untuk variabel yang berskala kategorik menggunakan metode ROCK dengan menetapkan

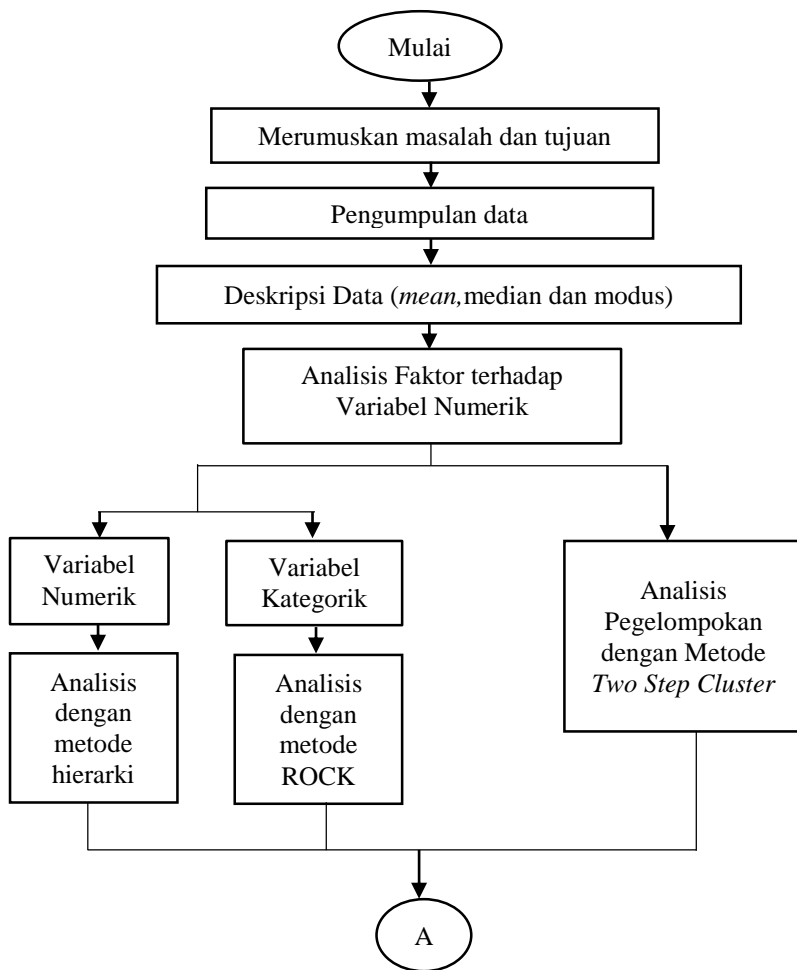
nilai *threshold* (θ) sebesar 0.05, 0.10, 0.15, 0.20, 0.25, 0.30, 0.35, 0.40, 0.45 dan 0.50.

5. Melakukan perhitungan rasio S_W dan S_B untuk menentukan nilai optimum untuk metode ROCK yang selanjutnya dinyatakan sebagai *output* cluster kategorik.
6. Melakukan pengelompokan berdasarkan metode ensemble ROCK dengan melakukan penggabungan *output* cluster numerik dan *output* cluster kategorik yang kemudian dikelompokkan menggunakan metode ROCK dengan menetapkan nilai *threshold* (θ) sebesar 0.05, 0.10, 0.15, 0.20, 0.25, 0.30, 0.35, 0.40, 0.45 dan 0.50.
7. Melakukan perhitungan rasio S_W dan S_B untuk menentukan jumlah cluster optimum berdasarkan data campuran sebagai hasil dari *final cluster*.
8. Melakukan pengelompokan variabel numerik dan kategorik menggunakan metode *Two Step Cluster*.
9. Melakukan perhitungan rasio S_W dan S_B dari hasil pengelompokan dengan metode *Two Step Cluster*.

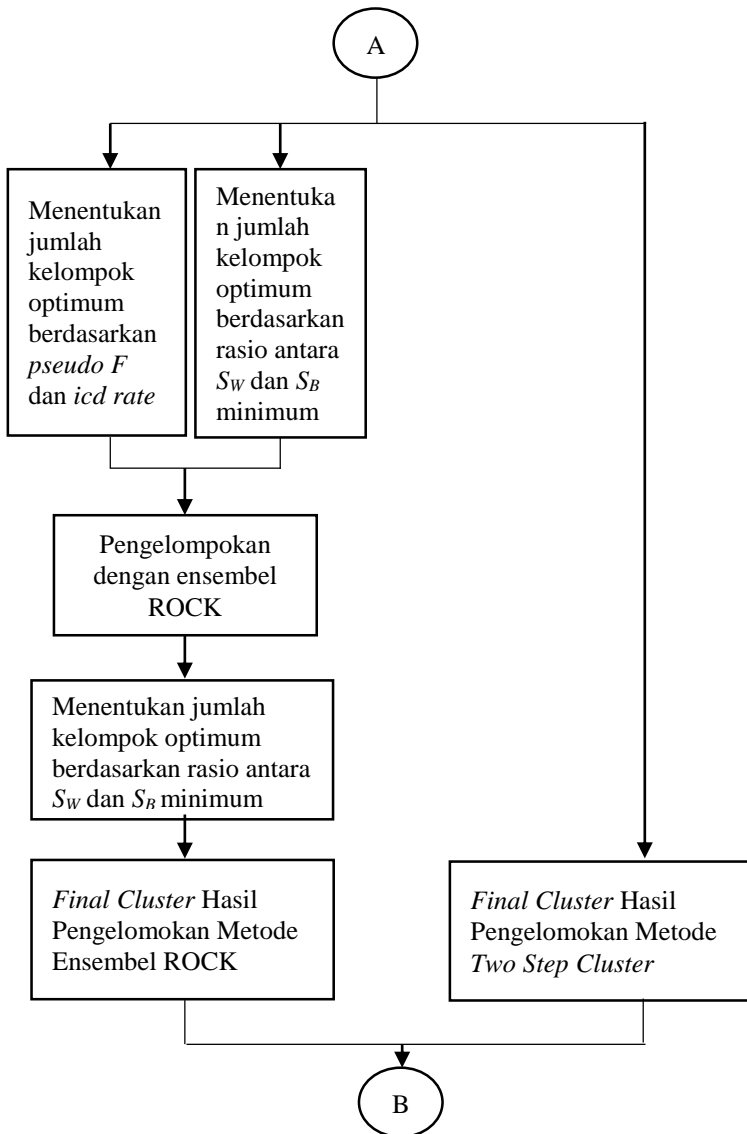
Serta untuk menyelesaikan tujuan keempat langkah analisisnya adalah

1. Membandingkan nilai rasio S_W dan S_B dari hasil pengelompokan dengan metode ensemble ROCK dan metode *Two Step Cluster*.
2. Melakukan pengujian perbedaan rata-rata antar kelompok dengan metode MANOVA untuk mengetahui apakah kelompok yang dihasilkan berbeda atau tidak.
3. Melakukan pengujian univariat untuk mengetahui variabel penyebab perbedaan antar kelompok dengan menggunakan ANOVA untuk variabel berskala numerik dan Uji Mann Whitney untuk variabel berskala kategorik.
4. Membuat kesimpulan dan saran.

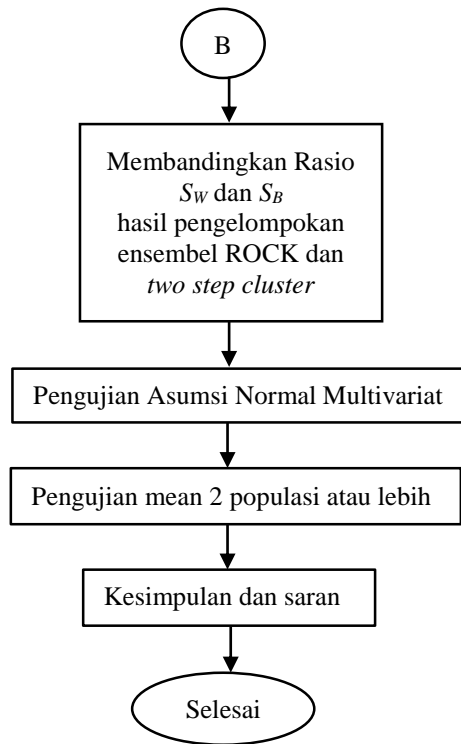
Langkah penelitian sebagaimana yang telah dijelaskan diatas dapat disajikan dalam diagram alir sebagai berikut



Gambar 3.1 Diagram Alir Penelitian



Gambar 3.1 Diagram Alir Penelitian (Lanjutan)



Gambar 3.1 Diagram Alir Penelitian (Lanjutan)

BAB IV ANALISIS DAN PEMBAHASAN

Kabupaten Bondowoso terdiri dari 23 kecamatan dengan 219 desa didalamnya. Pada penelitian ini dilakukan analisis terkait karakteristik dari desa-desa di Kabupaten Bondowoso untuk selanjutnya didapatkan hasil pengelompokan desa berdasarkan karakteristik-karakteristik dalam Indeks Pembangunan Desa.

4.1 Karakteristik Desa di Kabupaten Bondowoso

Karakteristik desa berdasarkan variabel numerik yang dimiliki dapat ditinjau dari segi sarana kesehatan, tempat jual beli, ketersediaan listrik, keuangan desa dan sarana pendidikan di tiap desa. Karakteristik desa di Bondowoso berdasarkan data numerik ditunjukkan dalam Tabel 4.1.

Tabel 4.1 Karakteristik Data Numerik

Variabel	Jumlah	Mean	Min	Max
Poskesdes (Unit)	152	-	0	1
T. Praktik Bidan (Unit)	220	-	0	4
T. Praktik Dokter (Unit)	72	-	0	10
Pasar Permanen (Unit)	21	-	0	1
Toko Kelontong (Unit)	8396	-	0	420
Keluarga Pengguna Listrik PLN (%)	-	96.32	0	100
PAD (Jutaan Rupiah)	-	43.33	0	386
ADD (Jutaan Rupiah)	-	101.83	0	170
SD/MI (Unit)	691	-	0	16
SMP/MTS (Unit)	215	-	0	6
Jarak terdekat menuju SMA/SMK/MA (km)	-	4.93	0	44.5

Dari segi sarana kesehatan terdapat Pos Kesehatan Desa (Poskesdes) sebagai tempat pelayanan kesehatan dasar bagi penduduk desa. Jumlah poskesdes yang ada di Kabupaten Bondowoso adalah 152 unit, dimana masing-masing terdapat satu unit poskesdes dalam tiap desa. Data ini menunjukkan masih terdapat 67 desa yang tidak memiliki poskesdes di wilayahnya.

Fasilitas kesehatan lainnya, untuk pemeriksaan kesehatan penduduk desa juga dapat memeriksa dirinya pada tempat praktik dokter yang terdapat pada wilayah-wilayah desanya. Tercatat terdapat 72 tempat praktik dokter yang tersebar di 219 desa/kelurahan di Kabupaten Bondowoso. Selain itu, untuk keperluan persalinan, rata-rata dalam setiap desa di Kabupaten Bondowoso telah memiliki 1 tempat praktik bidan di wilayahnya.

Untuk tempat jual beli barang, jumlah pasar permanen yang terdapat di Kabupaten Bondowoso masih sangat kurang. Hanya 21 desa yang memiliki pasar permanen di wilayahnya. Namun demikian, tersedia fasilitas lain untuk memenuhi kebutuhan akan barang tertentu bagi penduduk desa, yaitu telah tersedianya banyak warung atau toko kelontong yang terdapat dalam wilayah desa. Tercatat dalam setiap desa memiliki rata-rata 38 warung atau toko kelontong.

Dari segi keberadaan aliran listrik, hampir seluruh desa telah teraliri listrik oleh PLN, dimana rata-rata persentase keluarga dalam tiap desa yang rumahnya telah tersambung listrik sudah mencapai 96%. Adapun dari segi keuangan desa, rata-rata tiap desa di Kabupaten Bondowoso memperoleh anggaran dari pemerintah berupa Alokasi Dana Desa sebesar 101,8 Juta Rupiah, sedangkan untuk Pendapatan Asli Desa sendiri rata-rata sebesar 43.3 Juta Rupiah.

Dari segi sarana pendidikan pada umumnya dalam setiap desa masih terdapat SD/MI dan SMP/MTS di dalam wilayahnya, dimana rata-rata terdapat tiga SD/MI dan satu SMP/MTS dalam tiap desa. Sedangkan untuk sekolah menengah seperti SMA/SMK/MA karena jumlahnya memang dibatasi per kecamatan, maka banyak desa yang tidak memiliki SMA/SMK/MA dalam wilayahnya. Untuk itu setiap desa rata-rata menempuh jarak 4.9 km untuk menuju SMA/SMK/MA terdekat.

Karakteristik desa-desa di Kabupaten Bondowoso berdasarkan variabel-variabel berskala kategorik dapat ditunjukkan melalui modus atau kategori yang paling sering

muncul dalam setiap desa. Hal ini sebagaimana ditunjukkan dalam Tabel 4.2

Tabel 4.2 Karakteristik Data Kategorik

Variabel	Modus	Jumlah	Persentase
Kondisi Penerangan di Jalan Utama	5	143	65%
Bahan Bakar untuk Memasak	5	114	52%
Sumber Air untuk Minum	2	96	44%
Sumber Air untuk Mandi/Mencuci	2	82	37%
Fasilitas Buang Air Besar	0	113	52%
Kualitas Fasilitas Komunikasi Seluler	4	106	48%
Kelengkapan Pemerintahan Desa	5	217	99%
Aset/Kekayaan Desa	3	157	72%
Tingkat Pendidikan Kepala Desa	4	88	40%
Tingkat Pendidikan Sekretaris Desa	4	136	62%

Dari segi aspek kebutuhan dasar dapat dilihat melalui kondisi penerangan di jalan utama, bahan bakar untuk memasak, sumber air untuk minum, sumber air untuk mandi/mencuci, fasilitas buang air besar dan kualitas fasilitas komunikasi seluler. Variabel kondisi penerangan di jalan utama 65 % desa di Bondowoso berada dalam kategori 5, artinya telah ada penerangan di jalan utama desa dengan jenis penerangannya berupa listrik yang diusahakan oleh pemerintah. Variabel bahan bakar utama untuk memasak 52% desa di Bondowoso berada dalam kategori 5, artinya sebagian besar keluarga telah menggunakan LPG atau gas kota, dan ada pangkalan atau agen atau penjual LPG di wilayah desanya. Variabel sumber air untuk minum 44% desa di Bondowoso berada dalam kategori 2, artinya sumber air untuk minum sebagian besar keluarga berasal dari mata air. Begitupun untuk variabel sumber air untuk mandi/mencuci juga 37% desa di Kabupaten Bondowoso berada pada kategori 2, artinya sumber air untuk mandi/mencuci sebagian besar keluarga berasal dari mata air. Namun dari segi variabel fasilitas buang air besar masih 52% desa di Bondowoso

berada pada kategori 0, artinya desa yang fasilitas buang air besar sebagian besar keluarga adalah bukan jamban. Sedangkan untuk variabel fasilitas komunikasi seluler 48% desa di Bondowoso berada pada kategori 4, artinya desa yang tidak ada *Base Transceiver Station* (BTS), tetapi ada sinyal telepon seluler/*handphone* yang kuat.

Dari segi pemerintahan desa dapat ditinjau dari variabel kelengkapan pemerintahan desa, aset/kekayaan desa, tingkat pendidikan kepala desa dan tingkat pendidikan sekretaris desa. Variabel kelengkapan pemerintahan desa 99% desa di Bondowoso berada dalam kategori 5, artinya desa yang ada BPD dan ada kantor kepala desa di dalam wilayah desa. Variabel aset atau kekayaan desa 72% desa di Bondowoso berada dalam kategori 3, artinya desa mempunyai 3 (tiga) dari kelima jenis aset/pendapatan desa. Untuk variabel tingkat pendidikan Kepala Desa, 40% desa di Bondowoso berada pada kategori 4, artinya desa yang mempunyai kepala desa dengan pendidikan tertinggi yang ditamatkannya adalah SMU sederajat. Begitu juga dengan variabel tingkat pendidikan sekretaris desa, 62% desa di Bondowoso berada pada kategori 4, artinya desa yang mempunyai sekretaris desa dengan pendidikan tertinggi yang ditamatkannya adalah SMU sederajat.

4.2 Analisis Faktor pada Data Numerik

Analisis faktor dilakukan untuk mereduksi jumlah variabel numerik yang digunakan dalam pengelompokan. Sebelum dilakukan analisis faktor dilakukan pengujian dependensi menggunakan uji Bartlett. Uji Bartlett pada penelitian ini digunakan untuk mengetahui hubungan antara variabel-variabel numerik pada desa-desa di Kabupaten Bondowoso. Hasil pengujian KMO dan Bartlett menunjukkan *P-value* sebesar 0.000 sehingga lebih kecil dari α yaitu sebesar 0.05 (lampiran 7). Artinya dapat disimpulkan bahwa matriks korelasi 11 variabel tersebut bukan merupakan matriks identitas sehingga ada hubungan antara ke-11 variabel numerik pada desa-desa di Kabupaten Bondowoso. Berdasarkan nilai *eigen*, kemudian ditentukan jumlah komponen baru yang terbentuk. Sesuai dengan Tabel 4.3, maka terbentuk

empat komponen baru karena keempat komponen tersebut memiliki nilai *eigen* yang lebih besar dari 1 dimana dengan empat komponen baru yang terbentuk mampu menjelaskan persentase keragaman variansi sebesar 62.5 %.

Tabel 4.3 *Eigen Value* dan Persentase Variansi Kumulatif

Komponen	<i>Eigen Values</i>	%Variansi	% Variansi Kumulatif
1	2.77	25.19	25.19
2	1.57	14.26	39.44
3	1.51	13.76	53.20
4	1.03	9.33	62.53
5	0.89	8.05	70.59
6	0.83	7.53	78.11
7	0.67	6.10	84.21
8	0.53	4.85	89.07
9	0.46	4.19	93.26
10	0.45	4.08	97.34

Untuk mengetahui variabel-variabel yang digunakan masuk ke dalam kelompok komponen 1, 2, 3, atau 4 dengan memperhatikan nilai *loading factor*. Berikut nilai *loading factor* berdasarkan *rotate component matrix* untuk setiap variabel.

Tabel 4.4 Nilai *Loading Factor*

Variabel	PC1	PC2	PC3	PC4
Poskesdes	-0.16	0.32	0.01	-0.63
T. Praktik Bidan	0.58	0.31	0.26	-0.09
T. Praktik Dokter	0.74	-0.02	0.22	-0.33
Pasar Permanen	0.29	-0.05	0.51	-0.48
Toko Kelontong	0.83	0.06	0.01	0.30
Keluarga Pengguna Listrik PLN	0.05	0.88	0.00	-0.03
PAD	0.06	0.24	0.21	0.53
ADD	-0.38	-0.09	0.10	0.60
SD/MI	0.14	-0.01	0.75	0.24
SMP/MTS	0.06	0.22	0.77	0.03
Jarak terdekat menuju SMA/SMK/MA	-0.15	-0.87	-0.17	0.02

Berdasarkan Tabel 4.4, maka diketahui variabel yang masuk dalam komponen pertama adalah Tempat Praktik Bidan, Tempat Praktik Dokter dan Toko Kelontong. Variabel yang masuk dalam komponen kedua adalah Keluarga Pengguna Listrik PLN dan Jarak terdekat menuju SMA/SMK/MA. Variabel yang masuk kedalam komponen ketiga adalah Pasar Permanen, SD/MI dan SMP/MTS. Serta variabel yang masuk dalam komponen keempat adalah Poskesdes, PAD dan ADD.

Selanjutnya pada setiap komponen yang terbentuk dipilih variabel yang memiliki nilai *loading factor* maksimum absolut sebagai nilai yang mewakili komponen tersebut. Pada komponen pertama terpilih variabel Toko Kelontong, pada komponen kedua terpilih variabel keluarga pengguna listrik PLN, pada komponen ketiga terpilih variabel SMP/MTS dan pada komponen keempat terpilih variabel Poskesdes. Dengan demikian dari analisis faktor diperoleh 4 variabel numerik yaitu Toko Kelontong, keluarga pengguna listrik PLN, SMP/MTS dan Poskesdes yang selanjutnya akan digunakan dalam analisis pengelompokan.

4.3 Pengelompokan Data Numerik

Pengelompokan data numerik menggunakan teknik pengelompokan hierarki *agglomerative* dengan jarak *Euclidean* dan pautan *single linkage*, *complete linkage*, dan *average linkage*. Dalam setiap tautan dipilih jumlah cluster optimum berdasarkan nilai *Pseudo-F* maksimum dan kemudian dipilih metode tautan terbaik dengan nilai *ICD rate* minimum.

Pengelompokan desa di Kabupaten Bondowoso menggunakan tautan *single linkage* dengan jumlah dua, tiga, dan empat cluster diperoleh nilai *Pseudo-F* berturut-turut adalah sebesar 87.59, 139.03 dan 140.55. Dengan melihat nilai *Pseudo-F* maksimum pada tautan *single linkage* maka jumlah cluster optimum yang terbentuk adalah empat cluster. Sedangkan pengelompokan desa di Kabupaten Bondowoso menggunakan tautan *complete linkage* dengan jumlah dua, tiga, dan empat cluster diperoleh nilai *Pseudo-F* berturut-turut adalah sebesar 315.04, 191.66 dan 241.63 sehingga jumlah cluster optimum dengan

menggunakan pautan *complete linkage* adalah dua cluster. Adapun hasil pengelompokan desa di Kabupaten Bondowoso menggunakan tautan *average linkage* dengan jumlah dua, tiga, dan empat cluster diperoleh nilai *Pseudo-F* berturut-turut adalah sebesar 258.82, 139.03 dan 189.87 sehingga jumlah cluster optimum yang terbentuk dengan menggunakan tautan *average linkage* adalah dua cluster. Rincian dari nilai *Pseudo-F* dapat dilihat pada Tabel 4.5.

Tabel 4.5 Nilai *Pseudo-F* Pengelompokan Data Numerik

Pautan	Jumlah Cluster	<i>Pseudo-F</i>	<i>ICD Rate</i>
<i>Single Linkage</i>	2	87.59	0.71
	3	139.03	0.44
	4	140.55	0.34
<i>Complete Linkage</i>	2	315.04	0.41
	3	191.66	0.36
	4	241.63	0.23
<i>Average Linkage</i>	2	258.82	0.46
	3	139.03	0.44
	4	189.87	0.27

Selanjutnya dari ketiga metode pautan tersebut dibandingkan berdasarkan nilai *ICD Rate* sehingga diperoleh metode tautan terbaik untuk pengelompokan desa di Kabupaten Bondowoso berdasarkan data numerik. Metode tautan terbaik dalam mengelompokkan desa di Kabupaten Bondowoso adalah dengan menggunakan pautan *single linkage* dengan empat cluster, karena nilai *ICD Rate* dari pautan ini adalah minimum dibandingkan dengan nilai *ICD Rate* kedua pautan lainnya yaitu sebesar 0.34. Dengan demikian pengelompokan optimum desa di Kabupaten Bondowoso berdasarkan data numerik menggunakan pautan *single linkage* dengan jumlah cluster yang terbentuk adalah 4 cluster, yang terdiri dari 214 desa dalam cluster pertama, 2 desa dalam cluster kedua, 1 desa dalam cluster ketiga dan 2 desa dalam cluster keempat. Daftar anggota dari setiap cluster hasil pengelompokan numerik dapat dilihat dalam Lampiran 10.

4.4 Pengelompokan Data Kategorik

Pengelompokan data kategorik menggunakan metode ROCK. Dalam penelitian ini digunakan nilai *threshold* sebagai batas penentuan tetangga sebesar 0.05, 0.10, 0.15, 0.20, 0.25, 0.30, 0.35, 0.40, 0.45, dan 0.50 dimana jika jarak antar pengamatan lebih besar dari nilai *threshold* maka menunjukkan antar pengamatan tersebut bertetangga, untuk kemudian diukur nilai *link* antar pengamatan. Besarnya nilai *link* akan dipengaruhi oleh nilai *threshold* yang digunakan.

Dari sepuluh nilai *threshold* yang digunakan akan dipilih jumlah cluster optimum berdasarkan nilai rasio antara simpangan baku di dalam kelompok (S_W) dengan simpangan baku diantara kelompok (S_B). Semakin kecil nilai rasio yang diperoleh menunjukkan bahwa homogenitas yang maksimum di dalam kelompok serta heterogenitas yang maksimum antar kelompok.

Tabel 4.6 Rasio S_W dan S_B dalam Pengelompokan ROCK

<i>Threshold</i>	Jumlah Cluster	S_W	S_B	Rasio
0.05	3	1.25	3.00	0.42
0.10	3	0.67	5.71	0.12
0.15	2	0.85	16.65	0.05
0.20	3	0.93	13.51	0.07
0.25	2	1.25	10.56	0.12
0.30	3	0.74	8.30	0.09
0.35	2	0.31	4.29	0.07
0.40	2	0.31	4.05	0.08
0.45	3	0.54	3.00	0.18
0.50	1	0.00	-	-

Berdasarkan Tabel 4.6 maka nilai minimum rasio antara simpangan baku dalam kelompok dengan simpangan baku antar kelompok adalah 0.05. Oleh karena itu untuk pengelompokan desa di Kabupaten Bondowoso berdasarkan data kategorik menggunakan nilai *threshold* sebesar 0.15 dengan jumlah cluster optimum yang terbentuk adalah sebanyak 2 cluster Dimana sebanyak 72 desa masuk dalam cluster pertama dan 147 desa

lainnya masuk dalam cluster kedua. Daftar anggota dari setiap cluster hasil pengelompokan kategorik dapat dilihat dalam Lampiran 11.

4.5 Pengelompokan Data Campuran Menggunakan Metode Ensembl ROCK

Hasil pengelompokan data numerik dan pengelompokan data kategorik yang telah didapatkan dari pembahasan diatas dianggap sebagai variabel baru yang memiliki skala data kategorik. Hasil pengelompokan data numerik selanjutnya disebut sebagai *output* cluster numerik yang terdiri dari empat kategori dan hasil pengelompokan data kategorik disebut sebagai *output* cluster kategorik yang terdiri dari dua kategori. Tabel 4.7 menunjukkan struktur data baru dengan variabel berupa hasil cluster data numerik dan hasil cluster data kategorik.

Tabel 4.7 Struktur Data dengan Variabel Baru

No	Desa/Kelurahan	<i>Output Cluster</i> Numerik	<i>Output Cluster</i> Kategorik
1	Sucolor	1	1
2	Pujerbaru	1	1
3	Tanahwulan	1	1
⋮	⋮	⋮	⋮
217	Gunungsari	1	2
218	Sumberpandan	3	1
219	Pekauman	1	2

Karena hasil pengelompokan pada data numerik dan kategorik telah dianggap sebagai variabel baru dengan skala kategorik maka selanjutnya kembali dilakukan pengelompokan desa di Kabupaten Bondowoso dengan menggunakan metode ROCK. Nilai *threshold* yang digunakan tetap sebesar 0.05, 0.10, 0.15, 0.20, 0.25, 0.30, 0.35, 0.40, 0.45, dan 0.50, namun pada nilai *threshold* 0.35 sampai dengan 0.50 terjadi *error* karena tidak mampu mengelompokkan beberapa pengamatan sehingga penggunaan *threshold* 0.35 sampai dengan 0.50 tidak dilanjutkan dalam analisis selanjutnya. Dari ketiga nilai *threshold* yang

digunakan akan dipilih jumlah cluster optimum berdasarkan rasio dari S_W dengan S_B .

Tabel 4.8 Rasio S_W dan S_B Pengelompokan Ensemble ROCK

<i>Threshold</i>	Jumlah Cluster	S_W	S_B	Rasio
0.05	1	8.84×10^{-16}	-	-
0.10	1	8.84×10^{-16}	-	-
0.15	2	2.64×10^{-1}	0.90	0.29
0.20	3	5.76×10^{-1}	8.81	0.07
0.25	3	5.89×10^{-1}	9.90	0.06
0.30	2	1.32×10^{-1}	13.72	0.01

Berdasarkan Tabel 4.8 maka nilai minimum rasio antara simpangan baku dalam kelompok dengan simpangan baku antar kelompok adalah 0.01. Oleh karena itu untuk pengelompokan desa di Kabupaten Bondowoso berdasarkan data campuran numerik dan kategorik menggunakan nilai *threshold* sebesar 0.30 dengan jumlah cluster optimum yang terbentuk adalah sebanyak 2 cluster Dimana sebanyak 71 desa masuk dalam cluster pertama dan 148 desa masuk dalam cluster kedua. Daftar anggota dari setiap cluster hasil pengelompokan ensemble ROCK dapat dilihat dalam Lampiran 12.

4.6 Pengelompokan dengan Metode *Two Step Cluster*

Ada dua tahap dalam analisis *Two Step Cluster* yaitu tahap *preclustering* (Pembentukan *subcluster-subcluster* kecil) dan dilanjutkan dengan tahap pengelompokan hirarki pada setiap *subcluster*. Ukuran jarak yang digunakan adalah jarak *log-likelihood*, karena merupakan jarak berdasarkan probabilitas yang dapat menggabungkan antara variabel numerik dan kategorik. Proses pengelompokan dengan metode *two step cluster* sama dengan saat menggunakan metode ensemble ROCK, yaitu dilakukan pengelompokan terhadap data numerik menggunakan metode *two step cluster* untuk mendapatkan *output* cluster numerik, selanjutnya dilakukan pengelompokan terhadap variabel kategorik dengan metode *two step cluster* untuk mendapatkan *output* cluster

kategorik. Kemudian dari *output* cluster numerik dan kategorik ini yang selanjutnya dikelompokkan kembali dengan metode *two step cluster* untuk mendapatkan hasil *final cluster*. Berikut ini jumlah kelompok yang terbentuk dari hasil pengelompokan data numerik dan data kategorik menggunakan metode *two step cluster*.

Tabel 4.9 Hasil Pengelompokan Numerik dan Kategorik *Two Step Cluster*

Data	Jumlah Cluster	Anggota Cluster	Kualitas Pengelompokan
Variabel Numerik	3	1 = 141 desa 2 = 57 desa 3 = 21 desa	Baik
Variabel Kategorik	2	1 = 128 desa 2 = 91 desa	Buruk

Selanjutnya dilakukan pengelompokan kembali terhadap *output* cluster numerik dan *output* cluster kategorik dengan menggunakan metode *two step cluster* dan dilakukan perhitungan rasio simpangan baku dalam kelompok (S_w) dengan simpangan baku antara kelompok (S_B). Tabel 4.10 menunjukkan hasil *final cluster* menggunakan metode *two step cluster* dan hasil perhitungan rasio simpangan baku.

Tabel 4.10 *Final Cluster* Metode *Two Step Cluster*

Jumlah Cluster	Kualitas Pengelompokan	S_w	S_B	Rasio
6	Baik	1.09	5.16	0.21

Dengan demikian, pengelompokan desa di Kabupaten Bondowoso dengan menggunakan metode *two step cluster* menghasilkan cluster akhir sebanyak 6 cluster dengan nilai rasio simpangan baku dalam kelompok dengan simpangan baku antar kelompok sebesar 0.21. Anggota dari tiap cluster yaitu, cluster pertama terdiri dari 82 desa, cluster kedua terdiri dari 59 desa, cluster ketiga terdiri dari 36 desa, cluster keempat terdiri dari 21 desa, cluster kelima terdiri dari 11 desa dan cluster keenam terdiri dari 10 desa. Daftar anggota dari setiap cluster hasil pengelompokan *two step cluster* dapat dilihat dalam Lampiran 13.

4.7 Perbandingan Metode Ensemble ROCK dengan *Two Step Cluster*

Untuk dapat mengukur kebaikan evaluasi hasil pengelompokan desa di Kabupaten Bondowoso baik dengan menggunakan metode ensemble ROCK dan *Two Step Cluster* dapat menggunakan nilai rasio antara simpangan baku dalam kelompok (S_w) dengan simpangan baku antara kelompok (S_B). Semakin kecil nilai rasio yang diperoleh menunjukkan semakin baik hasil pengelompokan yang diperoleh, karena semakin maksimum kehomogenitasan didalam kelompok dan keheterogenitasan diantara kelompok. Tabel 4.11 menunjukkan perbandingan antara nilai rasio simpangan baku dalam kelompok dan diantara kelompok hasil dari *final cluster* menggunakan metode ensemble ROCK dan *two step cluster*.

Tabel 4.11 Perbandingan Rasio Ensemble ROCK dengan *Two Step Cluster*

Metode	Jumlah Cluster	Rasio
Ensemble ROCK	2	0.01
<i>Two Step Cluster</i>	6	0.21

Dari Tabel 4.11 nilai rasio dari pengelompokan dengan menggunakan metode ensemble ROCK lebih kecil dibandingkan dengan menggunakan metode *two step cluster* sehingga dapat disimpulkan bahwa hasil pengelompokan desa di Bondowoso dengan menggunakan metode ensemble ROCK lebih baik dibandingkan dengan menggunakan metode *two step cluster*.

4.8 Karakteristik Desa Hasil Pengelompokan Ensemble ROCK

Pengelompokan desa di Kabupaten Bondowoso dengan menggunakan metode ensemble ROCK telah menghasilkan 2 kelompok desa. Persebaran desa-desa yang masuk dalam kelompok pertama dan kelompok kedua disajikan dalam Gambar 4.1



Gambar 4.1 Persebaran Desa hasil *Final Cluster*

Karakteristik dari setiap kelompok desa digunakan untuk dapat mengidentifikasi desa yang tergolong daerah tertinggal. Karakteristik desa berdasarkan variabel numerik yang dimiliki dapat ditinjau dari segi sarana kesehatan, tempat jual beli, ketersediaan listrik, keuangan desa dan sarana pendidikan di tiap desa. Pada Tabel 4.12 dapat dilihat karakteristik variabel numerik antara desa pada cluster pertama dan cluster kedua.

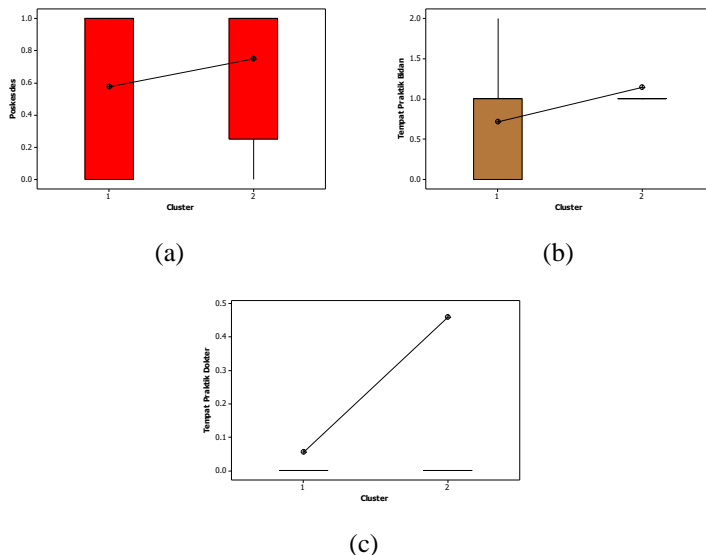
Tabel 4.12 Perbandingan Karakteristik Numerik Antar Cluster

Variabel	Cluster 1				Cluster 2			
	Sum	Mean	Min	Max	Sum	Mean	Min	Max
Poskesdes (Unit)	41	-	0	1	111	-	0	1
T. Praktik Bidan (Unit)	51	-	0	2	169	-	0	4
T. Praktik Dokter (Unit)	4	-	0	2	68	-	0	10
Pasar Permanen (Unit)	0	-	0	0	21	-	0	1
Toko Kelontong (Unit)	1956	-	0	125	6440	-	0	420

Tabel 4.12 Perbandingan Karakteristik Numerik Antar Cluster (Lanjutan)

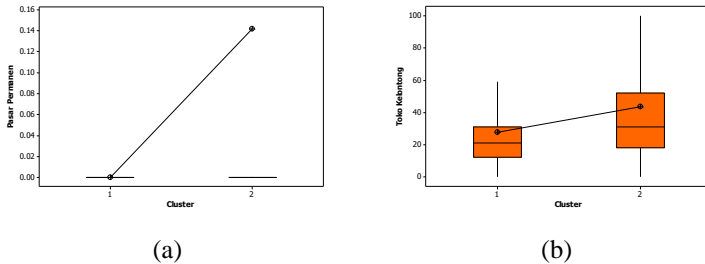
Variabel	Cluster 1				Cluster 2			
	<i>Sum</i>	<i>Mean</i>	<i>Min</i>	<i>Max</i>	<i>Sum</i>	<i>Mean</i>	<i>Min</i>	<i>Max</i>
Keluarga Pengguna Listrik PLN (%)	-	89.7	0	100	-	99.50	78	100
PAD (Jutaan Rupiah)	-	29.44	3	305	-	50	0	386
ADD (Jutaan Rupiah)	-	108.9	3	170	-	98.4	0	160
SD/MI (Unit)	205	-	1	7	486	-	0	16
SMP/MT S (Unit)	48	-	0	6	167	-	0	6
Jarak terdekat menuju SMA/SM K/MA (km)	-	9.1	0	44.5	-	2.91	0	12

Dari segi sarana kesehatan dasar, hanya 41 desa dari 71 desa di cluster 1 yang memiliki poskesdes di wilayahnya artinya, masih terdapat kira-kira 44% desa pada cluster 1 yang tidak memiliki fasilitas poskesdes. Nilai ini lebih sedikit dibandingkan pada desa dalam cluster 2, dimana terdapat 111 dari 148 desa telah memiliki poskesdes di wilayahnya. Artinya hanya 25% desa yang belum memiliki poskesdes pada cluster 2. Untuk sarana kesehatan lainnya ditunjang dengan adanya tempat praktik bidan dan tempat praktik dokter, dimana pada desa-desa dalam cluster 1 terdapat 51 tempat praktik bidan dan 4 tempat praktik dokter. Sedangkan pada desa dalam cluster 2 terdapat 169 tempat praktik bidan dan 68 tempat praktik dokter. Visualisasi perbedaan antara sarana kesehatan pada desa dalam cluster 1 dan 2 dapat dilihat pada Gambar 4.2.



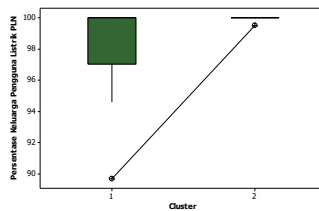
Gambar 4.2 Perbandingan Karakteristik Sarana Kesehatan , (a) Poskesdes, (b) Tempat Praktik Bidan, dan (c) Tempat Praktik Dokter.

Dari segi tempat untuk jual beli, perbedaan sangat terlihat karena pada desa dalam cluster 1 tidak terdapat sama sekali pasar permanen di wilayahnya. Hal ini berbanding terbalik dimana terdapat 21 pasar permanen di wilayah desa-desa pada cluster 2. Ini artinya seluruh pasar permanen yang ada di Kabupaten Bondowoso berada pada cluster 2. Namun demikian, masih banyak ditemui keberadaan dar toko kelontong baik dalam desa-desa pada cluster 1 dan 2, dimana rata-rata terdapat 27 toko kelontong pada desa di cluster 1 sedangkan pada desa di cluster 2 terdapat rata-rata 43 toko kelontong. Visualisasi perbedaan antara tempat jual beli barang pada desa dalam cluster 1 dan 2 dapat dilihat pada Gambar 4.3.



Gambar 4.3 Perbandingan Karakteristik (a) Pasar Permanen dan (b) Toko Kelontong

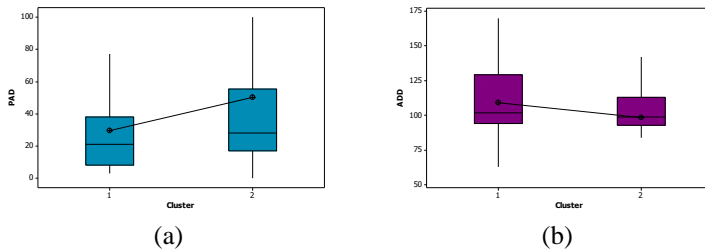
Dari segi ketersediaan arus listrik, persentase keluarga pengguna listrik PLN pada desa dalam cluster 1 ataupun 2 sudah cukup tinggi. Namun demikian, persentase keluarga pengguna listrik pada desa dalam cluster lebih kecil yaitu sebesar 89.7% jika dibandingkan dengan persentase keluarga pengguna listrik PLN pada desa dalam cluster 2 yaitu mencapai 99.5%. Visualisasi perbedaan antara persentase keluarga pengguna listrik pln pada desa dalam cluster 1 dan 2 dapat dilihat pada Gambar 4.4.



Gambar 4.4 Perbandingan Karakteristik Persentase Keluarga Pengguna Listrik PLN

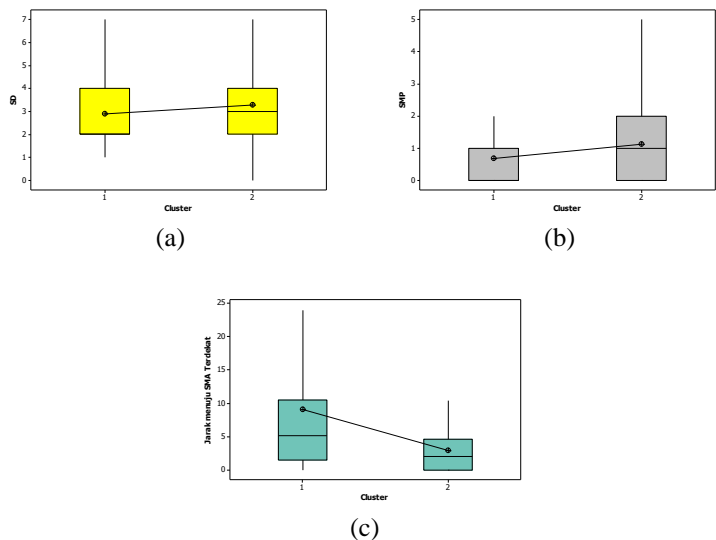
Dari segi keuangan desa, desa pada cluster 1 memiliki Pendapatan Asli Desa yang lebih kecil dibandingkan dengan desa pada cluster 2. Rata-rata desa pada cluster 1 memiliki Pendapatan Asli Desa sebesar 29.4 Juta Rupiah sedangkan desa pada cluster 2 memiliki rata-rata Pendapatan Asli Desa sebesar 50 Juta Rupiah. Disisi lain kondisi desa pada cluster 1 pun dapat dikatakan lebih rendah dbandingkan desa pada cluster 2, hal ini salah satunya dapat terlihat dari lebih besarnya rata-rata anggaran yang diperoleh dari

pemerintah berupa Alokasi Dana Desa, dimana pada desa dalam cluster 1 rata-rata memperoleh Alokasi Dana Desa sebesar 108.9 Juta Rupiah sedangkan desa pada cluster 2 memperoleh rata-rata sebesar 98.4 Juta Rupiah. Visualisasi perbedaan antara keuangan desa pada desa dalam cluster 1 dan 2 dapat dilihat pada Gambar 4.5.



Gambar 4.5 Perbandingan Karakteristik Keuangan Desa, (a) PAD dan (b) ADD

Dari segi sarana pendidikan jumlah sekolah dasar dan sekolah menengah pertama pada desa dalam cluster 1 jauh lebih sedikit dibandingkan pada desa dalam cluster 2. Pada desa dalam cluster 1 tercatat terdapat 205 SD/MI dan 49 SMP/MTS sedangkan pada desa dalam cluster 2 terdapat 486 SD/MI dan 167 SMP/MTS. Begitupun untuk jarak yang harus ditempuh untuk menuju SMA/SMK/MA terdekat, dimana untuk desa pada cluster 1 rata-rata harus menempuh jarak yang lebih jauh yaitu sebesar 9.1 km dibandingkan dengan desa pada cluster 2 yang hanya menempuh rata-rata 2.9 km untuk sampai pada SMA/SMK/MA terdekat. Visualisasi perbedaan antara sarana pendidikan pada desa dalam cluster 1 dan 2 dapat dilihat pada Gambar 4.6.



Gambar 4.6 Perbandingan Karakteristik Sarana Pendidikan, (a) SD/MI, (b) SMP/MTS, dan (c) Jarak Menuju SMA/SMK/MA terdekat

Karakteristik desa-desa di Kabupaten Bondowoso berdasarkan variabel-variabel berskala kategorik dapat ditunjukkan melalui modus atau kategori yang paling sering muncul dalam setiap desa. Perbandingan karakteristik desa berdasarkan variabel kategorik antara desa cluster 1 dengan cluster 2 disajikan pada Tabel 4.13.

Tabel 4.13 Perbandingan Karakteristik Variabel Kategorik

Variabel	Cluster 1			Cluster 2		
	Mode	Sum	Persentase	Mode	Sum	Persentase
Kondisi penerangan di Jalan utama	3	29	41%	5	121	82%
Bahan bakar untuk memasak	1	64	90%	5	110	74%

Tabel 4.13 Perbandingan Karakteristik Variabel Kategorik (Lanjutan)

Variabel	Cluster 1			Cluster 2		
	Mode	Sum	Persentase	Mode	Sum	Persentase
Sumber air untuk minum	2	60	85%	3	82	55%
Sumber air untuk mandi/mencuci	2	42	59%	3	59	40%
Fasilitas BAB	0	44	62%	0	69	47%
Kualitas fasilitas komunikasi seluler	2	30	42%	4	77	52%
Kelengkapan pemerintahan desa	5	70	99%	5	147	99%
Aset/keayaan desa	3	63	89%	3	94	64%
Tingkat pendidikan Kades	0	34	48%	4	64	43%
Tingkat pendidikan Sekdes	4	53	75%	4	83	56%

Dari 10 variabel kategorik dalam penelitian ini, 6 variabel menunjukkan perbedaan kategori mayoritas pada desa di cluster 1 dan cluster 2. Untuk variabel kondisi penerangan di jalan utama 41% desa di cluster 1 jenis penerangannya sudah berupa listrik namun diusahakan dari non pemerintah, sedangkan pada desa di cluster 2, 82% jenis penerangan sudah berupa listrik yang diupayakan oleh pemerintah. Untuk variabel bahan bakar utama untuk memasak, desa di cluster 1 90% penduduknya masih menggunakan kayu bakar sedangkan desa di cluster 2, 82% penduduknya telah menggunakan LPG. Untuk sumber air baik untuk minum maupun mandi/mencuci, desa di cluster 1 mayoritas

penduduknya masih bersumber dari sungai sedangkan desa di cluster 2 sudah mayoritas menggunakan air yang bersumber dari mata air. Untuk variabel fasilitas komunikasi seluler 42% desa di cluster 1 tidak memiliki sinyal seluler yang kuat sedangkan desa di cluster 2 52% desa didalamnya memiliki sinyal seluler yang kuat. Serta dari sisi SDM Pemerintah Desanya, 48% desa di cluster 1 tidak memiliki kepala desa sedangkan untuk desa di cluster 2 43% desa memiliki kepala desa dengan tingkat pendidikan terakhir merupakan tamatan SMA.

Dengan demikian, berdasarkan uraian mengenai karakteristik desa-desa yang terdapat dalam cluster satu dan dua menunjukkan adanya kecenderungan karakteristik desa pada cluster satu memiliki kondisi yang lebih rendah dibandingkan dengan karakteristik desa pada cluster dua. Oleh karena itu kelompok desa pada cluster satu diidentifikasi sebagai kelompok desa tertinggal yang ada di Kabupaten Bondowoso dimana terdiri dari 71 desa dari total keseluruhan 219 desa atau dengan kata lain persentase desa tertinggal di Kabupaten Bondowoso sebesar 32%.

4.9 Pengujian Beda Rata-rata Antar Kelompok

Setelah didapatkan hasil pengelompokan optimum dilakukan pengujian beda rata-rata antar kelompok. Pengujian dilakukan untuk mengetahui apakah hasil pengelompokan telah memberikan kelompok-kelompok yang berbeda secara signifikan. Secara multivariat pengujian dilakukan dengan menggunakan metode MANOVA. Oleh karena itu, dalam menguji perbedaan antar cluster hanya memperhatikan variabel yang berskala numerik. Namun untuk mengetahui variabel apa yang berbeda di antara cluster 1 dan cluster 2, variabel-variabel dalam penelitian ini dipandang sebagai data univariat, dimana untuk variabel numerik dilakukan pengujian menggunakan ANOVA *one way* dan untuk variabel kategorik dilakukan pengujian menggunakan uji Mann Whitney.

Sebelum dilakukan analisis dengan metode MANOVA, dilakukan dahulu pengujian normalitas mutivariat dari data. Pengujian normalitas multivariat menggunakan uji Mshapiro

Wilks. Hasil *mshapiro test* menunjukkan nilai *P-value* sebesar 2.2×10^{-16} sehingga lebih kecil dari nilai α yaitu sebesar 0.05 (lampiran 14). Oleh karena itu pengujian ini menunjukkan data tidak berdistribusi normal multivariat. Namun dalam penelitian ini, data yang digunakan dianggap telah berdistribusi normal multivariat, sehingga analisis dapat diteruskan untuk melakukan pengujian beda rata-rata antar kelompok.

Selain itu dilakukan juga pengujian dependensi untuk mengetahui hubungan antar variabel pada setiap cluster. Hasil pengujian KMO dan Bartlett menunjukkan *P-value* sebesar 0.000 sehingga lebih kecil dari α yaitu sebesar 0.05 (lampiran 7). Artinya dapat disimpulkan bahwa matriks korelasi 11 variabel tersebut bukan merupakan matriks identitas sehingga ada hubungan antara ke-11 variabel tersebut pada cluster satu dan dua

Pengujian beda rata-rata antar kelompok menggunakan MANOVA diperoleh hasil nilai statistik *Wilks lamda* sebesar 0.766 dengan nilai transformasi dalam *F* sebesar 5.755 dan *P-value* sebesar 0.000 (lampiran 15). Hal ini berarti bahwa nilai rata-rata dari setiap kelompok adalah berbeda secara multivariat. Oleh karena itu, dapat dinyatakan bahwa hasil pengelompokan dengan metode ensemble ROCK telah menghasilkan kelompok-kelompok yang berbeda secara signifikan.

Untuk mengetahui variabel numerik yang berbeda antara cluster 1 dengan cluster 2 dilakukan pengujian ANOVA *one way*. Tabel 4.14 menunjukkan hasil pengujian ANOVA terhadap 11 variabel numerik dalam penelitian ini.

Tabel 4.14 Hasil Pengujian ANOVA *One Way*

Variabel	<i>P-value</i>
Poskesdes	0.009
T. Praktik Bidan	0.000
T. Praktik Dokter	0.010
Pasar Permanen	0.001
Toko Kelontong	0.015
Keluarga Pengguna Listrik PLN	0.000
PAD	0.021

Tabel 4.14 Hasil Pengujian ANOVA *One Way* (Lanjutan)

Variabel	<i>P-value</i>
ADD	0.018
SD/MI	0.148
SMP/MTS	0.004
Jarak terdekat menuju SMA/SMK/MA	0.000

Berdasarkan Tabel 4.14, dengan menggunakan α sebesar 0.05 maka dari kesebelas variabel numerik yang digunakan, 10 variabel memiliki *p-value* yang lebih kecil dari α yang artinya 10 variabel tersebut signifikan berbeda antara desa di cluster 1 dan cluster 2. Adapun variabel yang tidak berbeda signifikan antara desa di cluster 1 dan cluster 2 adalah variabel jumlah SD/MI. Hal ini sesuai dengan yang ditunjukkan dalam Gambar 4.6(a) dimana terlihat hampir tidak perbedaan pada desa di cluster pertama dan cluster kedua.

Sedangkan untuk mengetahui variabel kategorik yang berbeda antara desa di cluster pertama dan cluster kedua menggunakan uji Mann Whitney. Tabel 4.15 menunjukkan hasil pengujian Mann Whitney terhadap 10 variabel kategorik.

Tabel 4.15 Hasil Pengujian Mann Whittney

Variabel	<i>P-value</i>
Kondisi Penerangan di Jalan Utama	0.000
Bahan Bakar untuk Memasak	0.000
Sumber Air untuk Minum	0.000
Sumber Air untuk Mandi/Mencuci	0.000
Fasilitas Buang Air Besar	0.002
Kualitas Fasilitas Komunikasi Seluler	0.000
Kelengkapan Pemerintahan Desa	0.930
Aset/Kekayaan Desa	0.071
Tingkat Pendidikan Kepala Desa	0.003
Tingkat Pendidikan Sekretaris Desa	0.109

Berdasarkan Tabel 4.15, dengan menggunakan α sebesar 0.05, maka dari 10 variabel kategorik yang digunakan dalam penelitian, tujuh variabel diantaranya memiliki *p-value* yang lebih

kecil dari α sehingga dapat disimpulkan tujuh variabel tersebut signifikan berbeda antara desa di cluster pertama dan cluster kedua. Adapun tiga variabel yang ternyata tidak berbeda signifikan antara desa di cluster pertama dan cluster kedua adalah variabel kelengkapan pemerintahan desa, aset/kekayaan desa dan tingkat pendidikan Sekretaris Desa.

(halaman ini sengaja dikosongkan)

BAB V

KESIMPULAN DAN SARAN

5.1 Kesimpulan

Kesimpulan yang didapatkan dari hasil penelitian ini adalah sebagai berikut.

1. Desa-desa di Kabupaten Bondowoso ditinjau dari segi sarana kesehatan, tempat jual beli, penggunaan listrik, keuangan desa, sarana pendidikan, sarana kebutuhan dasar, dan pemerintahan desa memiliki karakteristik yaitu
 - a. Dari segi sarana kesehatan 69% desa di Bondowoso memiliki poskesdes di wilayahnya sebagai sarana dasar kesehatan. Selain itu, sarana kesehatan lainnya ditunjang oleh adanya Tempat Praktik Bidan dan Tempat Praktik Dokter tersedia cukup banyak di beberapa wilayah desa.
 - b. Dari segi keberadaan pasar, jumlah pasar permanen yang ada di Kabupaten Bondowoso sangat terbatas yaitu sebanyak 21 unit, namun kondisi ini ditunjang dengan keberadaan toko kelontong yang jumlahnya rata-rata sebanyak 38 unit dalam setiap desa.
 - c. Dari segi penggunaan listrik, rata-rata persentase keluarga yang telah menggunakan listrik dari PLN dalam setiap desa di Bondowoso telah mencapai 96%.
 - d. Dari segi keuangan daerah rata-rata Pendapatan Asli Desa adalah sebesar 43.3 Juta Rupiah dan Alokasi Dana Desa adalah sebesar 101.8 Juta Rupiah.
 - e. Dari segi pendidikan rata-rata terdapat 3 SD/MI dan 1 SMP/MTS dalam setiap desa di Kabupaten Bondowoso. Sedangkan rata-rata jarak untuk menuju SMA/SMK/MA terdekat adalah sebesar 4.9 km.
 - f. Dari segi kondisi sarana kebutuhan dasar, mayoritas desa di Bondowoso telah menggunakan listrik sebagai sumber penerangan jalan utamanya, menggunakan LPG sebagai bahan bakar utama memasak, menggunakan air

yang bersumber dari mata air untuk keperluan minum, mandi/mencuci, serta sinyal seluler yang sudah kuat di wilayah desa-desa di Bondowoso. Namun untuk fasilitas BAB mayoritas desa masih belum menggunakan jamban.

- g. Dari segi pemerintahan desa, mayoritas desa di Kabupaten Bondowoso memiliki Kantor Desa dan BPD di wilayahnya, memiliki 3 dari 5 aset desa. Namun untuk tingkat pendidikan dari Kepala Desa dan Sekretaris Desa umumnya masih berupa tamatan SMA sederajat.
2. Hasil pengelompokan desa di Kabupaten Bondowoso menggunakan metode ensembel ROCK yaitu
 - a. Pengelompokan menggunakan data numerik diperoleh jumlah cluster optimum sebanyak 4 cluster dengan menggunakan pautan *single linkage*.
 - b. Pengelompokan menggunakan data kategorik diperoleh jumlah cluster optimum sebanyak 2 cluster dengan menggunakan *threshold* sebesar 0.15.
 - c. Pengelompokan menggunakan data campuran diperoleh jumlah cluster optimum sebanyak 2 cluster dengan menggunakan *threshold* 0.30. Anggota dari tiap cluster yaitu, cluster pertama terdiri dari 71 desa dan cluster kedua terdiri dari 148 desa.
3. Hasil pengelompokan desa di Kabupaten Bondowoso dengan menggunakan metode *two step cluster* adalah terdiri dari 6 cluster desa, dengan kualitas pengelompokan dalam kategori baik. Anggota dari tiap cluster yaitu, cluster pertama terdiri dari 82 desa, cluster kedua terdiri dari 59 desa, cluster ketiga terdiri dari 36 desa, cluster keempat terdiri dari 21 desa, cluster kelima terdiri dari 11 desa dan cluster keenam terdiri dari 10 desa..
4. Evaluasi kebaikan hasil pengelompokan menunjukkan metode ensembel ROCK menghasilkan pengelompokan desa yang lebih baik dibandingkan dengan metode *two step cluster* karena nilai rasio simpangan baku didalam

kelompok dan simpangan baku diantara kelompok metode ensemble ROCK (0.01) lebih kecil dari nilai rasio simpangan baku didalam kelompok dan simpangan baku diantara kelompok dengan metode *two step cluster* (0.21). Hasil ini menunjukkan pengelompokan optimum desa di Kabupaten Bondowoso menghasilkan 2 cluster yang terdiri dari 71 desa masuk dalam cluster pertama dan 148 desa masuk dalam cluster kedua. Berdasarkan karakteristik yang dimiliki pada desa di cluster pertama dan kedua menunjukkan kondisi desa di cluster pertama memiliki kondisi yang lebih rendah sehingga digolongkan sebagai kelompok desa tertinggal. Dengan demikian persentase desa tertinggal yang terdapat di Kabupaten Bondowoso adalah sebesar 32%.

5.2 Saran

Saran yang dapat diberikan berdasarkan hasil penelitian ini adalah dengan didapatkannya kelompok desa tertinggal perlu mendapatkan perhatian khusus dari Pemerintah Kabupaten Bondowoso untuk dapat memajukan desa tersebut. Terbatasnya fasilitas pasar permanen serta masih rendahnya kualitas SDM di desa dapat menjadi prioritas utama pemerintah dalam memajukan desa tertinggal.

Selain itu, untuk penelitian selanjutnya yang menggunakan metode ensemble ROCK dalam melakukan pengelompokan berdasarkan data campuran dengan ukuran dimensi yang cukup besar dapat menyederhanakan ukuran dimensi terlebih dahulu agar mendapatkan hasil pengelompokan yang lebih optimum serta dalam melakukan pengujian beda antar kelompok dapat mengembangkan metode baru yang juga memperhatikan data kategorik dalam penentuan uji beda antar kelompok yang terbentuk.

(halaman ini sengaja dikosongkan)

DAFTAR PUSTAKA

- Alvionita. (2017). *Metode Ensembel ROCK dan SWFM untuk Pengelompokan Data Campuran Numerik dan Kategorik pada Kasus Akses Jeruk*. Surabaya: Statistika, Institut Teknologi Sepuluh Nopember.
- Anderson, T., & Sclove, S. (1974). *Introductory Statistical Analysis*. Boston: Houghton Mifflin,.
- Bappenas, & BPS. (2015). *Indeks Pembangunan Desa*. Jakarta: Bappenas.
- Bunkers, M. J., & James, R. M. (1996). Definition of Climate Regions in the Northern Plains Using an Objective Cluster Modification Technique. *J. Climate*.
- Daniel, W. W. (1989). *Statistika Nonparametrik Terapan*. Alih Bahasa : Alex Tri Kandjono W. Jakarta: PT. Gramedia.
- Dewangan, R. R., Sharma, L., & Akasapu, A. K. (2010). Fuzzy Clustering Technique for Numerical and Categorical Dataset. *International Journal on Computer Science and Engineering*, hal 75-80.
- Dewi, A. (2012). *Metode Cluster Ensemble untuk Pengelompokan Desa Pedesaan di Provinsi Riau*. Surabaya: Statistika, Institut Teknologi Sepuluh Nopember.
- Dutta, M., Mahanta, A. K., & Arun, K. P. (2005). QROCK : A Quick Version of the ROCK Algorithm for Clustering of Categorical Data. *Proceeding of the 15th IEEE International Conference on Data Engineering*.
- Guha, S., Rastogi, R., & Shim, K. (2000). ROCK : A Robust Clustering Algorithm for Categorical Attributes. *Proceeding of the 15th International Conference on Data Engineering*.
- Hair, J. F., Black, W. C., Babin, J., & Anderson, E. (2009). *Multivariate Data Analysis (seventh ed.)*. New Jersey: Prentice Hall Inc.

- Halkidi, Batistakis, & Vizirgiannis. (2001). On Clustering Validation Techniques. *Journal of Intelligent Systems*, 17:2/3, hal 107-145.
- He, Z., Xu, X., & Deng, S. (2005a). A Cluster Ensemble Method For Clustering Categorical Data. *Informaton Fusion*, hal 143-151.
- He, Z., Xu, X., & Deng, S. (2005b). Clustering Mixed Numeric and Categorical Data: A Cluster Ensemble Approach. *Departemen of Computer Science and Engineering , Harbin Institute of Technology*.
- Johnson, R. A., & Wicherin, D. W. (2007). *Applied Multivariate Statistical Analysis (sixth ed.)*. New Jersey: Pearson Education Inc.
- Kementerian DPD TT. (2016). *Petunjuk Pelaksanaan Identifikasi Masalah-masalah Ketertinggalan Kabupaten Daerah Tertinggal*. Jakarta: Direktorat Jenderal Pembangunan Daerah Tertinggal.
- Sharma, S. (1996). *Applied Multivariate Techniques*. New York: John Wiley and Sons Inc.
- Suguna, J., & Selvi, M. A. (2012). Ensemble Fuzzy Clustering for Mixed for Numerical and Categorical Data. *Intenational Journal of Computer Application*, 42, 19.
- Tyagi, A., & Sharma, S. (2012). Implementation of ROCK Clustering Algorithm for the Optimazation of Query Searching Time. *International Journal on Computer Science and Engineering*, Vol 4, No 05.
- Walpole, R. E. (1993). *Pengantar Statistika Edisi ke-3*. Alih Bahasa : Bambang Soemantri. Jakarta: PT Gramedia Pustaka Utama.

LAMPIRAN

Lampiran 1 Data Penelitian Variabel Numerik

No	Desa	X ₁	X ₂	X ₃	X ₄	X ₅	X ₆	X ₇	X ₈	X ₉	X ₁₀	X ₁₁
1	SUCOLOR	0	1	0	0	23	100	54	108	4	2	0
2	PUJERBARU	1	1	0	0	0	100	54	100	4	1	3.5
3	TANAHWULAN	1	1	0	0	21	100	42	100	3	0	4
4	MAESAN	1	1	3	0	15	100	49	92	1	0	2
5	GAMBANGAN	1	1	0	1	22	100	46	100	3	2	1.5
6	SUGER LOR	1	1	0	0	47	100	17	97	3	0	5
7	SUMBERPAKEM	0	1	0	0	54	100	19	95	2	0	5
8	SUMBERSARI	1	1	2	0	2	100	112	104	5	6	3
9	SUMBERANYAR	0	1	0	0	100	100	25	79	2	0	4
10	SUCOLOR	0	1	0	0	23	100	54	108	4	2	0
11	PENANGGUNGAN	1	1	1	0	32	100	63	96	1	1	0.7
12	PAKUNIRAN	1	2	0	0	28	100	24	104	2	2	0
13	GUNUNGSARI	1	1	0	0	40	100	3	106	2	0	0.2
14	SUMBERPANDAN	0	1	0	0	27	100	21	93	2	0	0
15	PEKAUMAN	1	1	0	0	31	100	169	96	2	1	0
16	WANISODO	1	0	0	0	12	100	29	91	1	0	1
17	DAWUHAN	1	1	0	0	56	100	80	94	2	1	2
18	KABUARAN	1	1	0	0	39	100	30	92	2	1	5
19	WONOSARI	0	1	0	0	59	100	63	100	4	0	5
20	DADAPAN	1	1	0	1	38	100	28	91	4	0	0
...
201	SEMPOL	1	1	0	0	45	100	16	123	4	1	4.3
202	PRAJEKAN LOR	1	1	0	0	61	100	323	134	2	0	2.8
203	PRAJEKAN KIDUL	1	3	7	1	43	100	13	113	4	3	0
204	CANGKRING	0	0	0	0	52	91	5	140	2	0	2
205	WALIDONO	0	0	0	0	58	100	32	102	4	0	2.3
206	SULING KULON	1	1	1	0	25	100	35	97	2	1	4
207	SOLOR	0	1	0	0	6	53	46	131	5	0	20
208	KLADI	0	1	0	0	16	100	33	122	2	0	18
209	BERCAK	1	2	0	0	32	100	56	146	3	2	5
210	CERMEE	0	2	2	1	300	100	235	117	6	0	10
211	SULING WETAN	1	1	0	0	14	100	10	135	3	1	8
212	RAMBAN WETAN	1	1	0	0	91	100	34	105	6	1	0
213	RAMBAN KULON	1	1	0	0	100	100	286	103	4	1	0
214	GRUJUGAN	0	1	0	0	56	100	15	94	3	1	0
215	BAJURAN	0	1	1	0	25	100	28	104	4	2	3
216	JIREK MAS	0	1	0	0	11	92	22	126	5	0	24
217	BATUSALANG	0	1	1	0	10	86	42	137	1	0	21
218	PELALANGAN	1	1	0	0	15	89	24	140	3	0	7
219	BATU AMPAR	0	1	0	0	26	32	37	126	2	0	21

Lampiran 2 Data Penelitian Variabel Kategorik

No	Desa	X ₁₂	X ₁₃	X ₁₄	X ₁₅	X ₁₆	X ₁₇	X ₁₈	X ₁₉	X ₂₀	X ₂₁
1	SUCOLOR	0	3	1	2	1	5	2	5	3	4
2	PUJERBARU	3.5	0	1	2	1	2	4	5	3	0
3	TANAHWULAN	4	5	1	2	2	2	4	5	3	0
4	MAESAN	2	3	5	3	2	5	4	5	3	3
5	GAMBANGAN	1.5	5	5	4	4	5	5	5	4	0
6	SUGER LOR	5	5	5	3	2	0	4	5	3	4
7	SUMBERPAKEM	5	5	5	3	1	0	4	5	4	4
8	SUMBERSARI	3	3	1	3	3	2	5	5	3	0
9	SUMBERANYAR	4	0	1	2	1	0	4	5	4	3
10	SUCOLOR	0.7	5	5	2	2	0	4	5	4	4
11	PENANGGUNGAN	0	3	1	3	1	0	5	5	3	0
12	PAKUNIRAN	0.2	3	5	2	1	0	4	5	3	5
13	GUNUNGSARI	0	5	1	4	4	0	4	5	3	0
14	SUMBERPANDAN	0	3	5	3	3	0	4	5	4	0
15	PEKAUMAN	1	3	1	2	2	0	4	5	3	0
16	WANISODO	2	3	5	2	2	5	4	5	3	4
17	DAWUHAN	5	3	1	2	2	3	5	5	3	0
18	KABUARAN	5	3	1	2	2	0	3	5	3	4
19	WONOSARI	0	5	5	4	4	5	4	5	5	4
20	DADAPAN	1	5	5	3	3	5	5	5	4	4
...	...										
201	SEMPOL	3	5	3	3	5	2	5	4	4	4
202	PRAJEKAN LOR	5	5	3	3	5	5	5	3	3	4
203	PRAJEKAN KIDUL	5	5	5	5	5	5	5	5	3	5
204	CANGKRING	5	5	3	3	0	4	5	3	0	4
205	WALIDONO	3	5	3	3	0	2	5	3	4	4
206	SULING KULON	3	5	3	3	3	4	5	3	4	4
207	SOLOR	3	1	2	1	0	2	5	3	3	3
208	KLADI	5	1	2	1	0	4	5	3	4	4
209	BERCAK	5	1	3	3	0	4	5	3	4	4
210	CERMEE	5	5	4	4	5	5	5	5	5	5
211	SULING WETAN	3	5	3	3	5	4	5	3	4	5
212	RAMBAN WETAN	3	1	3	3	0	4	5	4	5	4
213	RAMBAN KULON	5	5	4	4	0	4	5	4	4	5
214	GRUJUGAN	5	1	3	3	2	4	5	3	5	4
215	BAJURAN	3	1	2	2	2	2	5	3	5	5
216	JIREK MAS	0	1	3	3	0	2	5	3	4	4
217	BATUSALANG	5	1	2	1	0	4	5	3	4	4
218	PELALANGAN	3	5	3	3	0	2	5	3	5	4
219	BATU AMPAR	5	1	2	1	0	2	5	3	3	4

Lampiran 3 Syntax Pengelompokan Data Numerik

```
#Pautan single linkage
#Pautan single linkage
kluster.single = function(data,k){
  datanumerik = data.frame(data$X1, data$X5, data$X6,
  data$X10)
  d = dist(datanumerik, method = "euclidean")
  fit.sin = hclust(d, method = "single")
  single = cutree(fit.sin, k=k)
  hasil.kelompok.single = data.frame(data$Desa, single)
  print(fit.sin)
  print(single)
  print(hasil.kelompok.single)
}

#Pautan complete linkage
kluster.complete = function(data,k){
  datanumerik = data.frame(data$X1, data$X5, data$X6,
  data$X10)
  d = dist(datanumerik, method = "euclidean")
  fit.sin = hclust(d, method = "complete")
  complete = cutree(fit.sin, k=k)
  hasil.kelompok.complete = data.frame(data$Desa, complete)
  print(fit.sin)
  print(complete)
  print(hasil.kelompok.complete)
}

# pautan average linkage
kluster.average = function(data,k){
  datanumerik = data.frame(data$X1, data$X5, data$X6,
  data$X10)
  d = dist(datanumerik, method = "euclidean")
  fit.sin = hclust(d, method = "average")
  average = cutree(fit.sin, k=k)
```

Lampiran 3 Syntax Pengelompokan Data Numerik (Lanjutan)

```
hasil.kelompok.average = data.frame(data$Desa, average)
print(fit.sin)
print(average)
print(hasil.kelompok.average)
}
```


Lampiran 4 Syntax Pengelompokan Data Kategorik

```

datakategorik1
c(data$X12,data$X13,,data$X14,data$X15,data$X16,data$X1
7,data$X18,data$X19,data$X20,data$X21)
datakategorik2 = matrix (datakategorik1,219,10)
x = as.dummy(datakategorik2)
rc.05 = rockCluster(x, n=5, theta=0.05, debug=FALSE)
rc.10 = rockCluster(x, n=5, theta=0.10, debug=FALSE)
rc.15 = rockCluster(x, n=5, theta=0.15, debug=FALSE)
rc.20 = rockCluster(x, n=5, theta=0.20, debug=FALSE)
rc.25 = rockCluster(x, n=5, theta=0.25, debug=FALSE)
rc.30 = rockCluster(x, n=5, theta=0.30, debug=FALSE)
rc.35 = rockCluster(x, n=5, theta=0.35, debug=FALSE)
rc.40 = rockCluster(x, n=5, theta=0.40, debug=FALSE)
rc.45 = rockCluster(x, n=5, theta=0.45, debug=FALSE)
rc.50 = rockCluster(x, n=5, theta=0.50, debug=FALSE)
rf.05 = fitted(rc.05)
rf.10 = fitted(rc.10)
rf.15 = fitted(rc.15)
rf.20 = fitted(rc.20)
rf.25 = fitted(rc.25)
rf.30 = fitted(rc.30)
rf.35 = fitted(rc.35)
rf.40 = fitted(rc.40)
rf.45 = fitted(rc.45)
rf.50 = fitted(rc.50)
theta.05 = rf.05$cl
theta.10 = rf.10$cl
theta.15 = rf.15$cl
theta.20 = rf.20$cl
theta.25 = rf.25$cl
theta.30 = rf.30$cl
theta.35 = rf.35$cl
theta.40 = rf.40$cl
theta.45 = rf.45$cl

```

Lampiran 4 Syntax Pengelompokan Data Kategorik (Lanjutan)

```

theta.50 = rf.50$cl
Kelompok1 = data.frame(data$Desa, theta.05, theta.10,
theta.15, theta.20, theta.25)
Kelompok2 = data.frame(data$Desa, theta.30, theta.35,
theta.40, theta.45, theta.50)
model1<-read.table("E:/Tugas Akhir Dhany/Revisi
TA/p05.csv",sep="," ,header=TRUE)
model2<-read.table("E:/Tugas Akhir Dhany/Revisi
TA/p10.csv",sep="," ,header=TRUE)
model3<-read.table("E:/Tugas Akhir Dhany/Revisi
TA/p15.csv",sep="," ,header=TRUE)
model4<-read.table("E:/Tugas Akhir Dhany/Revisi
TA/p20.csv",sep="," ,header=TRUE)
model5<-read.table("E:/Tugas Akhir Dhany/Revisi
TA/p25.csv",sep="," ,header=TRUE)
model6<-read.table("E:/Tugas Akhir Dhany/Revisi
TA/p30.csv",sep="," ,header=TRUE)
model7<-read.table("E:/Tugas Akhir Dhany/Revisi
TA/p35.csv",sep="," ,header=TRUE)
model8<-read.table("E:/Tugas Akhir Dhany/Revisi
TA/p40.csv",sep="," ,header=TRUE)
model9<-read.table("E:/Tugas Akhir Dhany/Revisi
TA/p45.csv",sep="," ,header=TRUE)
model10<-read.table("E:/Tugas Akhir Dhany/Revisi
TA/p50.csv",sep="," ,header=TRUE)
mylogit.05 =
aov(theta.05~X12+X13+X14+X15+X16+X17+X18+X19+X2
0+X21, data=model1)
mylogit.10 =
aov(theta.10~X12+X13+X14+X15+X16+X17+X18+X19+X2
0+X21,data=model2)
mylogit.15 =
aov(theta.15~X12+X13+X14+X15+X16+X17+X18+X19+X2
0+X21, data=model3)

```

Lampiran 4 Syntax Pengelompokan Data Kategorik (Lanjutan)

```

mylogit.20 =
aov(theta.20~X12+X13+X14+X15+X16+X17+X18+X19+X20+X21, data=model4)
mylogit.25 =
aov(theta.25~X12+X13+X14+X15+X16+X17+X18+X19+X20+X21, data=model5)
mylogit.30 =
aov(theta.30~X12+X13+X14+X15+X16+X17+X18+X19+X20+X21, data=model6)
mylogit.35 =
aov(theta.35~X12+X13+X14+X15+X16+X17+X18+X19+X20+X21, data=model7)
mylogit.40 =
aov(theta.40~X12+X13+X14+X15+X16+X17+X18+X19+X20+X21, data=model8)
mylogit.45 =
aov(theta.45~X12+X13+X14+X15+X16+X17+X18+X19+X20+X21, data=model9)
mylogit.50 =
aov(theta.50~X12+X13+X14+X15+X16+X17+X18+X19+X20+X21, data=model10)
SSW.05 = sum((mylogit.05$residuals)^2)
SSW.10 = sum((mylogit.10$residuals)^2)
SSW.15 = sum((mylogit.15$residuals)^2)
SSW.20 = sum((mylogit.20$residuals)^2)
SSW.25 = sum((mylogit.25$residuals)^2)
SSW.30 = sum((mylogit.30$residuals)^2)
SSW.35 = sum((mylogit.35$residuals)^2)
SSW.40 = sum((mylogit.40$residuals)^2)
SSW.45 = sum((mylogit.45$residuals)^2)
SSW.50 = sum((mylogit.50$residuals)^2)
summary.05 = c(summary(mylogit.05))
summary.10 = c(summary(mylogit.10))
summary.15 = c(summary(mylogit.15))

```

Lampiran 4 Syntax Pengelompokan Data Kategorik (Lanjutan)

```

summary.20 = c(summary(mylogit.20))
summary.25 = c(summary(mylogit.25))
summary.30 = c(summary(mylogit.30))
summary.35 = c(summary(mylogit.35))
summary.40 = c(summary(mylogit.40))
summary.45 = c(summary(mylogit.45))
summary.50 = c(summary(mylogit.50))
SSB.05 = 17.96
SSB.10 = 65.27
SSB.15 = 277.25
SSB.20 = 365.02
SSB.25 = 111.48
SSB.30 = 137.78
SSB.35 = 18.37
SSB.40 = 16.44
SSB.45 = 20.53
SSB.50 = 0.00
SW.05=sqrt(SSW.05/(219-3))
SW.10=sqrt(SSW.10/(219-3))
SW.15=sqrt(SSW.15/(219-2))
SW.20 = sqrt(SSW.20/(219-3))
SW.25=sqrt(SSW.25/(219-2))
SW.30 = sqrt(SSW.30/(219-3))
SW.35=sqrt(SSW.35/(219-2))
SW.40 = sqrt(SSW.40/(219-2))
SW.45=sqrt(SSW.45/(219-3))
SW.50=sqrt(SSW.50/(219-1))
SW = rbind(SW.05, SW.10, SW.15, SW.20, SW.25, SW.30,
SW.35, SW.40, SW.45, SW.50)
SB.05 = sqrt(SSB.05/(3-1))
SB.10 = sqrt(SSB.10/(3-1))
SB.15 = sqrt(SSB.15/(2-1))
SB.20 = sqrt(SSB.20/(3-1))
SB.25 = sqrt(SSB.25/(2-1))

```

Lampiran 4 Syntax Pengelompokan Data Kategorik (Lanjutan)

```

SB.30 = sqrt(SSB.30/(3-1))
SB.35 = sqrt(SSB.35/(2-1))
SB.40 = sqrt(SSB.40/(2-1))
SB.45 = sqrt(SSB.05/(3-1))
SB.50 = sqrt(SSB.50/(1-1))
SB = rbind(SB.05, SB.10, SB.15, SB.20, SB.25, SB.30, SB.35,
SB.40, SB.45, SB.50)
Ratio.05 = SW.05/SB.05
Ratio.10 = SW.10/SB.10
Ratio.15 = SW.15/SB.15
Ratio.20 = SW.20/SB.20
Ratio.25 = SW.25/SB.25
Ratio.30 = SW.30/SB.30
Ratio.35 = SW.35/SB.35
Ratio.40 = SW.40/SB.40
Ratio.45 = SW.45/SB.45
Ratio.50 = SW.50/SB.50
Ratio.kategorik = rbind(Ratio.05, Ratio.10, Ratio.15, Ratio.20,
Ratio.25, Ratio.30, Ratio.35, Ratio.40, Ratio.45, Ratio.50))

```

Lampiran 5 Syntax Pengelompokan Ensembel ROCK

```

datakategorik1 = c(data$Numerik,data$Kategorik)
datakategorik2 = matrix (datakategorik1,219,2)
x = as.dummy(datakategorik2)
rc.05 = rockCluster(x, n=5, theta=0.05, debug=FALSE)
rc.10 = rockCluster(x, n=5, theta=0.10, debug=FALSE)
rc.15 = rockCluster(x, n=5, theta=0.15, debug=FALSE)
rc.20 = rockCluster(x, n=5, theta=0.20, debug=FALSE)
rc.25 = rockCluster(x, n=5, theta=0.25, debug=FALSE)
rc.30 = rockCluster(x, n=5, theta=0.30, debug=FALSE)
rc.35 = rockCluster(x, n=5, theta=0.35, debug=FALSE)
rc.40 = rockCluster(x, n=5, theta=0.40, debug=FALSE)
rc.45 = rockCluster(x, n=5, theta=0.45, debug=FALSE)
rc.50 = rockCluster(x, n=5, theta=0.50, debug=FALSE)
rf.05 = fitted(rc.05)
rf.10 = fitted(rc.10)
rf.15 = fitted(rc.15)
rf.20 = fitted(rc.20)
rf.25 = fitted(rc.25)
rf.30 = fitted(rc.30)
rf.35 = fitted(rc.35)
rf.40 = fitted(rc.40)
rf.45 = fitted(rc.45)
rf.50 = fitted(rc.50)
theta.05 = rf.05$cl
theta.10 = rf.10$cl
theta.15 = rf.15$cl
theta.20 = rf.20$cl
theta.25 = rf.25$cl
theta.30 = rf.30$cl
theta.35 = rf.35$cl
theta.40 = rf.40$cl
theta.45 = rf.45$cl
theta.50 = rf.50$cl

```

Lampiran 5 Syntax Pengelompokan Ensembel ROCK (Lanjutan)

```

Kelompok1 = data.frame(data$Desa, theta.05, theta.10,
theta.15, theta.20, theta.25)
Kelompok2 = data.frame(data$Desa, theta.30, theta.35,
theta.40, theta.45, theta.50)
model1<-read.table("E:/Tugas Akhir Dhany/Revisi
TA/d05.csv",sep="," ,header=TRUE)
model2<-read.table("E:/Tugas Akhir Dhany/Revisi
TA/d10.csv",sep="," ,header=TRUE)
model3<-read.table("E:/Tugas Akhir Dhany/Revisi
TA/d15.csv",sep="," ,header=TRUE)
model4<-read.table("E:/Tugas Akhir Dhany/Revisi
TA/d20.csv",sep="," ,header=TRUE)
model5<-read.table("E:/Tugas Akhir Dhany/Revisi
TA/d25.csv",sep="," ,header=TRUE)
model6<-read.table("E:/Tugas Akhir Dhany/Revisi
TA/d30.csv",sep="," ,header=TRUE)
mylogit.05 = aov(theta.05~Numerik+Kategorik, data=model1)
mylogit.10 = aov(theta.10~Numerik+Kategorik, data=model2)
mylogit.15 = aov(theta.15~Numerik+Kategorik, data=model3)
mylogit.20 = aov(theta.20~Numerik+Kategorik, data=model4)
mylogit.25 = aov(theta.25~Numerik+Kategorik, data=model5)
mylogit.30 = aov(theta.30~Numerik+Kategorik, data=model6)
SSW.05 = sum((mylogit.05$residuals)^2)
SSW.10 = sum((mylogit.10$residuals)^2)
SSW.15 = sum((mylogit.15$residuals)^2)
SSW.20 = sum((mylogit.20$residuals)^2)
SSW.25 = sum((mylogit.25$residuals)^2)
SSW.30 = sum((mylogit.30$residuals)^2)
summary.05 = c(summary(mylogit.05))
summary.10 = c(summary(mylogit.10))
summary.15 = c(summary(mylogit.15))
summary.20 = c(summary(mylogit.20))
summary.25 = c(summary(mylogit.25))
summary.30 = c(summary(mylogit.30))

```

Lampiran 5 Syntax Pengelompokan Ensembel ROCK (Lanjutan)

```

SSB.05 = 0.00
SSB.10 = 0.00
SSB.15 = 0.813
SSB.20 = 155.276
SSB.25 = 196.182
SSB.30 = 188.149
SW.05=sqrt(SSW.05/(219-1))
SW.10=sqrt(SSW.10/(219-1))
SW.15=sqrt(SSW.15/(219-2))
SW.20 = sqrt(SSW.20/(219-3))
SW.25=sqrt(SSW.25/(219-3))
SW.30 = sqrt(SSW.30/(219-2))
SW = rbind(SW.05, SW.10, SW.15, SW.20, SW.25, SW.30)
SB.05 = sqrt(SSB.05/(1-1))
SB.10 = sqrt(SSB.10/(1-1))
SB.15 = sqrt(SSB.15/(2-1))
SB.20 = sqrt(SSB.20/(3-1))
SB.25 = sqrt(SSB.25/(3-1))
SB.30 = sqrt(SSB.30/(2-1))
SB = rbind(SB.05, SB.10, SB.15, SB.20, SB.25, SB.30)
Ratio.05 = SW.05/SB.05
Ratio.10 = SW.10/SB.10
Ratio.15 = SW.15/SB.15
Ratio.20 = SW.20/SB.20
Ratio.25 = SW.25/SB.25
Ratio.30 = SW.30/SB.30
Ratio.kategorik = rbind(Ratio.05, Ratio.10, Ratio.15, Ratio.20,
Ratio.25, Ratio.30)

```


Lampiran 6 Output Statistika Deskriptif Seluruh Data**#Data Numerik**

Variable	Mean	Sum	Minimum	Maximum
X1	0.6941	152.0000	0.0000	1.0000
X2	1.0046	220.0000	0.0000	4.0000
X3	0.3288	72.0000	0.0000	10.0000
X4	0.0959	21.0000	0.0000	1.0000
X5	38.34	8396.00	0.00	420.00
X6	96.32	21093.49	0.00	100.00
X7	43.33	9490.00	0.00	386.00
X8	101.83	22300.00	0.00	170.00
X9	3.155	691.000	0.000	16.000
X10	0.9817	215.0000	0.0000	6.0000
X11	4.917	1076.900	0.000	44.500

#Data Kategorik

Variable	Mode	N for Mode
X12	5	143
X13	5	114
X14	2	96
X15	2	82
X16	0	113
X17	4	106
X18	5	217
X19	3	157
X20	4	88

Lampiran 7 Pemeriksaan KMO dan Uji Bartlett

Kaiser-Meyer-Olkin Measure of Sampling Adequacy.		.662
Bartlett's Test of Sphericity	Approx. Chi-Square	476.226
	df	55
	Sig.	.000

Lampiran 8 Persentase Variansi Kumulatif

Component	Initial Eigenvalues		
	Total	% of Variance	Cumulative %
1	2.77	25.19	25.19
2	1.57	14.26	39.44
3	1.51	13.76	53.20
4	1.03	9.33	62.53
5	0.89	8.05	70.59
6	0.83	7.53	78.11
7	0.67	6.10	84.21
8	0.53	4.85	89.07
9	0.46	4.19	93.26
10	0.45	4.08	97.34
11	0.29	2.66	100.00

Lampiran 9 *Rotated Component Matrix*

	Component			
	1	2	3	4
X1	-0.16	0.32	0.01	-0.63
X2	0.58	0.31	0.26	-0.09
X3	0.74	-0.02	0.22	-0.33
X4	0.29	-0.05	0.51	-0.48
X5	0.83	0.06	0.01	0.30
X6	0.05	0.88	0.00	-0.03
X7	0.06	0.24	0.21	0.53
X8	-0.38	-0.09	0.10	0.60
X9	0.14	-0.01	0.75	0.24
X10	0.06	0.22	0.77	0.03
X11	-0.15	-0.87	-0.17	0.02

Lampiran 10 Output Pengelompokan Data Numerik

No	Desa	Cluster	No	Desa	Cluster
1	Sucolor	1	31	Jambeanom	1
2	Pujerbaru	1	32	Jambesari	1
3	Tanahwulan	1	33	Pejagan	1
4	Maesan	1	34	Grujugan Lor	1
5	Gambangan	1	35	Tegalpasir	1
6	Suger Lor	1	36	Pengarang	1
7	Sumberpakem	1	37	Sumber Anyar	1
8	Sumbersari	1	38	Sukokerto	1
9	Sumberanyar	1	39	Sukowono	1
10	Penanggungan	1	40	Maskuning Wetan	1
11	Pakuniran	1	41	Maskuning Kulon	1
12	Gunungsari	1	42	Alassumur	1
13	Sumberpandan	1	43	Mengok	1
14	Pekauman	1	44	Kejayan	1
15	Wanisodo	1	45	Mangli	1
16	Dawuhan	1	46	Randucangkring	1
17	Kabuaran	1	47	Padasan	1
18	Wonosari	1	48	Sukodono	1
19	Dadapan	1	49	Pakisan	1
20	Taman	1	50	Tlogosari	1
21	Tegalmijin	1	51	Kembang	1
22	Grujugan Kidul	1	52	Gunosari	1
23	Kejawen	1	53	Trotosari	1
24	Sukosari	1	54	Sulek	1
25	Karangmelok	1	55	Jebung Kidul	1
26	Mengen	1	56	Jebung Lor	1
27	Kemirian	1	57	Patemon	1
28	Tamanan	1	58	Brambang Darussalam	1
29	Wonosuko	1	59	Pecalongan	1
30	Kalianyar	1	60	Nogosari	1

Lampiran 10 Output Pengelompokan Data Numerik (Lanjutan)

No	Desa	Cluster	No	Desa	Cluster
61	Sumberkemuning	1	91	Sukosari Lor	1
62	Sumberanom	1	92	Kerang	1
63	Sumberjeruk	1	93	Sukosari Kidul	1
64	Pucang Anom	1	94	Tegal Jati	1
65	Rejo Agung	2	95	Sumbersalam	1
66	Sukorejo	1	96	Pekalangan	1
67	Sumber Gading	1	97	Kasemek	1
68	Sumberwringin	1	98	Lojajar	1
69	Wonokusumo	1	99	Kajar	1
70	Mangli Wetan	1	100	Bataan	1
71	Gununganyar	1	101	Gebang	1
72	Jurangsapi	1	102	Dawuhan	1
73	Cindogo	1	103	Tenggarang	1
74	Kalitapen	1	104	Tangsil Kulon	1
75	Mrawan	1	105	Koncer Darul Aman	1
76	Taal	1	106	Pancoran	1
77	Tapen	1	107	Sukowiryo	1
78	Lombok Kulon	1	108	Kembang	1
79	Lombok Wetan	1	109	Nangkaan	1
80	Tumpeng	1	110	Tamansari	1
81	Jumpong	1	111	Dabasah	2
82	Tangsil Wetan	1	112	Badean	4
83	Pasarejo	1	113	Kotakulon	1
84	Bendoarum	1	114	Blindungan	1
85	Kapuran	1	115	Kademangan	1
86	Sumberkalong	1	116	Pejaten	1
87	Traktakan	1	117	Jetis	1
88	Wonosari	1	118	Pakuwesi	1
89	Pelalangan	1	119	Kupang	1
90	Koncer Kidul	3	120	Petung	1

Lampiran 10 Output Pengelompokan Data Numerik (Lanjutan)

No	Desa	Cluster	No	Desa	Cluster
121	Penambangan	1	151	Banyuwulu	1
122	Curahpoh	1	152	Jatitamban	1
123	Curahdami	1	153	Jatisari	1
124	Poncogati	1	154	Glingseran	1
125	Sumbersuko	1	155	Banyuputih	1
126	Silolembu	1	156	Wringin	1
127	Locare	1	157	Sumbercantin g	1
128	Sumber Salak	1	158	Karanganyar	1
129	Gadingsari	1	159	Sekarputih	1
130	Binakal	1	160	Tegalampel	1
131	Sumber Waru	1	161	Mandiro	1
132	Kembangan	1	162	Tanggulangun	1
133	Baratan	1	163	Klabang	1
134	Jeruk Soksok	1	164	Klabang Agung	1
135	Bendelan	1	165	Purnama	1
136	Sumber Tengah	1	166	Trebungan	1
137	Andungsari	1	167	Paguan	1
138	K U P A N G	1	168	Sumberkokap	1
139	Ardisaeng	1	169	Taman	1
140	Sumberdumpyo ng	1	170	Gentong	1
141	Gadingsari	1	171	Kemuningan	1
142	P A K E M	1	172	Kretek	1
143	Patemon	1	173	Karanganyar	1
144	P E T U N G	1	174	Blimbing	1
145	Ambulu	1	175	Karang Sengon	1
146	Bukor	1	176	Wonokerto	1
147	Sumbermalang	1	177	Klabang	1
148	Jambewungu	1	178	Besuk	1
149	GUBRIH	1	179	Klampokan	1
150	AMPELAN	1	180	Sumbersuko	1

Lampiran 10 Output Pengelompokan Data Numerik (Lanjutan)

No	Desa	Cluster	No	Desa	Cluster
181	Pandak	1	211	Ramban Wetan	1
182	Leprak	1	212	Ramban Kulon	1
183	Wonobojo	1	213	Grujugan	1
184	Sempol	1	214	Bajuran	1
185	Kalianyar	1	215	Jirek Mas	1
186	Kalisat	1	216	Batusalang	1
187	Jampit	1	217	Pelalangan	1
188	Sumber Rejo	1	218	Batu Ampar	1
189	Kaligedang	1	219	Bercak Asri	1
190	Sumbercanting	1			
191	Lanas	1			
192	Penang	1			
193	Klekean	1			
194	Gayam	1			
195	Lumutan	1			
196	Botolinggo	1			
197	Gayam Lor	1			
198	Tarum	1			
199	Bandilan	1			
200	Sempol	1			
201	Prajejan Lor	1			
202	Prajejan Kidul	1			
203	Cangkring	1			
204	Walidono	1			
205	Suling Kulon	1			
206	Solor	1			
207	Kladi	1			
208	Bercak	1			
209	Cermee	4			
210	Suling Wetan	1			

Lampiran 11 Output Pengelompokan Data Kategorik

No	Desa	Cluster	No	Desa	Cluster
1	Sucolor	1	31	Sumberkemuning	2
2	Pujerbaru	1	32	Sumberanom	2
3	Tanahwulan	1	33	Sumberjeruk	2
4	Maesan	2	34	Pucang Anom	2
5	Gambangan	2	35	Jambeanom	2
6	Suger Lor	2	36	Jambesari	2
7	Sumberpakem	2	37	Pejagan	2
8	Sumbersari	1	38	Grujugan Lor	2
9	Sumberanyar	1	39	Tegalpasir	2
10	Penanggungan	2	40	Pengarang	2
11	Pakuniran	1	41	Sumber Anyar	2
12	Gunungsari	1	42	Sukokerto	2
13	Sumberpandan	1	43	Sukowono	2
14	Pekauman	2	44	Maskuning Wetan	2
15	Wanisodo	1	45	Maskuning Kulon	2
16	Dawuhan	2	46	Alasumur	2
17	Kabuaran	1	47	Mengok	2
18	Wonosari	1	48	Kejayan	2
19	Dadapan	2	49	Mangli	2
20	Taman	2	50	Randucangkring	2
21	Tegalmijin	2	51	Padasan	1
22	Grujugan Kidul	2	52	Sukodono	2
23	Kejawan	2	53	Pakisan	2
24	Sukosari	2	54	Tlogosari	2
25	Karangmelok	2	55	Kembang	2
26	Mengen	2	56	Gunosari	1
27	Kemirian	2	57	Trotosari	1
28	Tamanan	2	58	Sulek	2
29	Wonosuko	2	59	Jebung Kidul	2
30	Kalianyar	2	60	Jebung Lor	2

Lampiran 11 Output Pengelompokan Data Kategori (Lanjutan)

No	Desa	Cluster	No	Desa	Cluster
61	Patemon	2	91	Traktakan	2
62	Brambang Darussalam	1	92	Wonosari	2
63	Pecalongan	1	93	Pelalangan	2
64	Nogosari	2	94	Koncer Kidul	2
65	Sukosari Lor	2	95	Sumbersalam	2
66	Kerang	2	96	Pekalangan	2
67	Sukosari Kidul	1	97	Kasemek	2
68	Tegal Jati	1	98	Lojajar	2
69	Rejo Agung	1	99	Kajar	2
70	Sukorejo	1	100	Bataan	2
71	Sumber Gading	2	101	Gebang	2
72	Sumberwringin	2	102	Dawuhan	1
73	Wonokusumo	2	103	Tenggarang	2
74	Mangli Wetan	2	104	Tangsil Kulon	2
75	Gununganyar	2	105	Koncer Darul Aman	2
76	Jurangsapi	2	106	Pancoran	2
77	Cindogo	2	107	Sukowiryo	2
78	Kalitapen	2	108	Kembang	2
79	Mrawan	2	109	Nangkaan	2
80	Taal	2	110	Tamansari	2
81	Tapen	2	111	Dabasah	2
82	Lombok Kulon	2	112	Badean	2
83	Lombok Wetan	1	113	Kotakulon	2
84	Tumpeng	2	114	Blindungan	2
85	Jumpong	2	115	Kademangan	2
86	Tangsil Wetan	2	116	Pejaten	2
87	Pasarejo	2	117	Jetis	2
88	Bendoarum	2	118	Pakuwesi	2
89	Kapuran	2	119	Kupang	2
90	Sumberkalong	2	120	Petung	2

Lampiran 11 Output Pengelompokan Data Kategori (Lanjutan)

No	Desa	Cluster	No	Desa	Cluster
121	Penambangan	2	151	Banyuwulu	1
122	Curahpoh	1	152	Jatitamban	2
123	Curahdami	2	153	Jatisari	2
124	Poncogati	2	154	Glingseran	1
125	Sumbersuko	2	155	Banyuputih	1
126	Silolembu	2	156	Wringin	2
127	Locare	2	157	Sumbercantin g	2
128	Sumber Salak	1	158	Karanganyar	2
129	Gadingsari	1	159	Sekarputih	2
130	Binakal	1	160	Tegalampel	2
131	Sumber Waru	1	161	Mandiro	1
132	Kembangan	1	162	Tanggulangun	1
133	Baratan	1	163	Klabang	1
134	Jeruk Soksok	1	164	Klabang Agung	1
135	Bendelan	1	165	Purnama	1
136	Sumber Tengah	2	166	Trebungan	1
137	Andungsari	1	167	Paguan	1
138	K U P A N G	2	168	Sumberkokap	1
139	Ardisaeng	1	169	Taman	1
140	Sumberdumpyo ng	2	170	Gentong	1
141	Gadingsari	1	171	Kemuningan	1
142	P A K E M	2	172	Kretek	1
143	Patemon	2	173	Karanganyar	2
144	P E T U N G	2	174	Blimbing	2
145	Ambulu	1	175	Karang Sengon	2
146	Bukor	2	176	Wonokerto	2
147	Sumbermalang	2	177	Klabang	2
148	Jambewungu	1	178	Besuk	2
149	Gubrih	1	179	Klampokan	2
150	Ampelan	1	180	Sumbersuko	2

Lampiran 11 Output Pengelompokan Data Kategori (Lanjutan)

No	Desa	Cluster	No	Desa	Cluster
181	Pandak	2	211	Ramban Wetan	2
182	Leprak	2	212	Ramban Kulon	2
183	Wonobojo	1	213	Grujugan	2
184	Sempol	1	214	Bajuran	1
185	Kalianyar	1	215	Jirek Mas	1
186	Kalisat	1	216	Batusalang	1
187	Jampit	1	217	Pelalangan	2
188	Sumber Rejo	1	218	Batu Ampar	1
189	Kaligedang	1	219	Bercak Asri	2
190	Sumbercanting	1			
191	Lanas	1			
192	Penang	1			
193	Klekean	1			
194	Gayam	2			
195	Lumutan	2			
196	Botolinggo	2			
197	Gayam Lor	2			
198	Tarum	1			
199	Bandilan	1			
200	Sempol	2			
201	Prajejan Lor	2			
202	Prajejan Kidul	2			
203	Cangkring	2			
204	Walidono	2			
205	Suling Kulon	2			
206	Solor	1			
207	Kladi	1			
208	Bercak	2			
209	Cermee	2			
210	Suling Wetan	2			

Lampiran 12 Output Pengelompokan Ensembel ROCK

No	Desa	Cluster	No	Desa	Cluster
1	Sucolor	1	31	Sumberkemuning	2
2	Pujerbaru	1	32	Sumberanom	2
3	Tanahwulan	1	33	Sumberjeruk	2
4	Maesan	2	34	Pucang Anom	2
5	Gambangan	2	35	Jambeanom	2
6	Suger Lor	2	36	Jambesari	2
7	Sumberpakem	2	37	Pejagan	2
8	Sumbersari	1	38	Grujugan Lor	2
9	Sumberanyar	1	39	Tegalpasir	2
10	Penanggungan	2	40	Pengarang	2
11	Pakuniran	1	41	Sumber Anyar	2
12	Gunungsari	1	42	Sukokerto	2
13	Sumberpandan	1	43	Sukowono	2
14	Pekauman	2	44	Maskuning Wetan	2
15	Wanisodo	1	45	Maskuning Kulon	2
16	Dawuhan	2	46	Alassumur	2
17	Kabuaran	1	47	Mengok	2
18	Wonosari	1	48	Kejayan	2
19	Dadapan	2	49	Mangli	2
20	Taman	2	50	Randucangkring	2
21	Tegalmijin	2	51	Padasan	1
22	Grujugan Kidul	2	52	Sukodono	2
23	Kejawen	2	53	Pakisan	2
24	Sukosari	2	54	Tlogosari	2
25	Karangmelok	2	55	Kembang	2
26	Mengen	2	56	Gunosari	1
27	Kemirian	2	57	Trotosari	1
28	Tamanan	2	58	Sulek	2
29	Wonosuko	2	59	Jebung Kidul	2
30	Kalianyar	2	60	Jebung Lor	2

Lampiran 12 Output Pengelompokan Ensembel ROCK (Lanjutan)

No	Desa	Cluster	No	Desa	Cluster
61	Patemon	2	91	Traktakan	2
62	Brambang Darussalam	1	92	Wonosari	2
63	Pecalongan	1	93	Pelalangan	2
64	Nogosari	2	94	Koncer Kidul	2
65	Sukosari Lor	2	95	Sumbersalam	2
66	Kerang	2	96	Pekalangan	2
67	Sukosari Kidul	1	97	Kasemek	2
68	Tegal Jati	1	98	Lojajar	2
69	Rejo Agung	2	99	Kajar	2
70	Sukorejo	1	100	Bataan	2
71	Sumber Gading	2	101	Gebang	2
72	Sumberwringin	2	102	Dawuhan	1
73	Wonokusumo	2	103	Tenggarang	2
74	Mangli Wetan	2	104	Tangsil Kulon	2
75	Gununganyar	2	105	Koncer Darul Aman	2
76	Jurangsapi	2	106	Pancoran	2
77	Cindogo	2	107	Sukowiryo	2
78	Kalitapen	2	108	Kembang	2
79	Mrawan	2	109	Nangkaan	2
80	Taal	2	110	Tamansari	2
81	Tapen	2	111	Dabasah	2
82	Lombok Kulon	2	112	Badean	2
83	Lombok Wetan	1	113	Kotakulon	2
84	Tumpeng	2	114	Blindungan	2
85	Jumpong	2	115	Kademangan	2
86	Tangsil Wetan	2	116	Pejaten	2
87	Pasarejo	2	117	Jetis	2
88	Bendoarum	2	118	Pakuwesi	2
89	Kapuran	2	119	Kupang	2
90	Sumberkalong	2	120	Petung	2

Lampiran 12 Output Pengelompokan Ensembel ROCK (Lanjutan)

No	Desa	Cluster	No	Desa	Cluster
121	Penambangan	2	151	Banyuwulu	1
122	Curahpoh	1	152	Jatitamban	2
123	Curahdami	2	153	Jatisari	2
124	Poncogati	2	154	Glingseran	1
125	Sumbersuko	2	155	Banyuputih	1
126	Silolembu	2	156	Wringin	2
127	Locare	2	157	Sumbercantin g	2
128	Sumber Salak	1	158	Karanganyar	2
129	Gadingsari	1	159	Sekarputih	2
130	Binakal	1	160	Tegalampel	2
131	Sumber Waru	1	161	Mandiro	1
132	Kembangan	1	162	Tanggulangun	1
133	Baratan	1	163	Klabang	1
134	Jeruk Soksok	1	164	Klabang Agung	1
135	Bendelan	1	165	Purnama	1
136	Sumber Tengah	2	166	Trebungan	1
137	Andungsari	1	167	Paguan	1
138	K U P A N G	2	168	Sumberkokap	1
139	Ardisaeng	1	169	Taman	1
140	Sumberdumpyo ng	2	170	Gentong	1
141	Gadingsari	1	171	Kemuningan	1
142	P A K E M	2	172	Kretek	1
143	Patemon	2	173	Karanganyar	2
144	P E T U N G	2	174	Blimbing	2
145	Ambulu	1	175	Karang Sengon	2
146	Bukor	2	176	Wonokerto	2
147	Sumbermalang	2	177	Klabang	2
148	Jambewungu	1	178	Besuk	2
149	Gubrih	1	179	Klampokan	2
150	Ampelan	1	180	Sumbersuko	2

Lampiran 12 Output Pengelompokan Ensembel ROCK (Lanjutan)

No	Desa	Cluster	No	Desa	Cluster
181	Pandak	2	211	Ramban Wetan	2
182	Leprak	2	212	Ramban Kulon	2
183	Wonoboyo	1	213	Grujugan	2
184	Sempol	1	214	Bajuran	1
185	Kalianyar	1	215	Jirek Mas	1
186	Kalisat	1	216	Batusalang	1
187	Jampit	1	217	Pelalangan	2
188	Sumber Rejo	1	218	Batu Ampar	1
189	Kaligedang	1	219	Bercak Asri	2
190	Sumbercanting	1			
191	Lanas	1			
192	Penang	1			
193	Klekean	1			
194	Gayam	2			
195	Lumutan	2			
196	Botolinggo	2			
197	Gayam Lor	2			
198	Tarum	1			
199	Bandilan	1			
200	Sempol	2			
201	Prajekean Lor	2			
202	Prajekean Kidul	2			
203	Cangkring	2			
204	Walidono	2			
205	Suling Kulon	2			
206	Solor	1			
207	Kladi	1			
208	Bercak	2			
209	Cermee	2			
210	Suling Wetan	2			

Lampiran 13 Output Pengelompokan *Two Step Cluster*

No	Desa	Cluster	No	Desa	Cluster
1	Sucolor	1	31	Sumberkemuning	3
2	Pujerbaru	2	32	Sumberanom	3
3	Tanahwulan	2	33	Sumberjeruk	2
4	Maesan	3	34	Pucang Anom	1
5	Gambangan	3	35	Jambeanom	2
6	Suger Lor	2	36	Jambesari	3
7	Sumberpakem	1	37	Pejagan	2
8	Sumbersari	4	38	Grujugan Lor	3
9	Sumberanyar	1	39	Tegalpasir	2
10	Penanggungan	2	40	Pengarang	3
11	Pakuniran	2	41	Sumber Anyar	4
12	Gunungsari	2	42	Sukokerto	2
13	Sumberpandan	5	43	Sukowono	2
14	Pekauman	3	44	Maskuning Wetan	1
15	Wanisodo	2	45	Maskuning Kulon	2
16	Dawuhan	2	46	Alassumur	2
17	Kabuaran	2	47	Mengok	2
18	Wonosari	1	48	Kejayan	1
19	Dadapan	3	49	Mangli	5
20	Taman	3	50	Randucangkring	2
21	Tegalmijin	5	51	Padasan	2
22	Grujugan Kidul	5	52	Sukodono	2
23	Kejawan	3	53	Pakisan	2
24	Sukosari	3	54	Tlogosari	1
25	Karangmelok	5	55	Kembang	2
26	Mengen	3	56	Gunosari	1
27	Kemirian	2	57	Trotosari	1
28	Tamanan	3	58	Sulek	2
29	Wonosuko	3	59	Jebung Kidul	2
30	Kalianyar	1	60	Jebung Lor	3

Lampiran 13 Output Pengelompokan *Two Step Cluster* (Lanjutan)

No	Desa	Cluster	No	Desa	Cluster
61	Patemon	2	91	Traktakan	3
62	Brambang Darussalam	3	92	Wonosari	3
63	Pecalongan	2	93	Pelalangan	3
64	Nogosari	1	94	Koncer Kidul	2
65	Sukosari Lor	1	95	Sumbersalam	4
66	Kerang	3	96	Pekalangan	3
67	Sukosari Kidul	2	97	Kasemek	2
68	Tegal Jati	2	98	Lojajar	2
69	Rejo Agung	1	99	Kajar	5
70	Sukorejo	6	100	Bataan	2
71	Sumber Gading	6	101	Gebang	4
72	Sumberwringin	2	102	Dawuhan	3
73	Wonokusumo	2	103	Tenggarang	2
74	Mangli Wetan	5	104	Tangsil Kulon	5
75	Gununganyar	5	105	Koncer Darul Aman	1
76	Jurangsapi	3	106	Pancoran	3
77	Cindogo	2	107	Sukowiryo	5
78	Kalitapen	3	108	Kembang	3
79	Mrawan	3	109	Nangkaan	5
80	Taal	2	110	Tamansari	3
81	Tapen	2	111	Dabasah	3
82	Lombok Kulon	3	112	Badean	4
83	Lombok Wetan	1	113	Kotakulon	4
84	Tumpang	2	114	Blindungan	3
85	Jumpong	2	115	Kademangan	5
86	Tangsil Wetan	2	116	Pejaten	4
87	Pasarejo	1	117	Jetis	3
88	Bendoarum	2	118	Pakuwesi	4
89	Kapuran	5	119	Kupang	2
90	Sumberkalong	2	120	Petung	2

Lampiran 13 Output Pengelompokan *Two Step Cluster* (Lanjutan)

No	Desa	Cluster	No	Desa	Cluster
121	Penambangan	5	151	Banyuwulu	1
122	Curahpoh	3	152	Jatitamban	6
123	Curahdami	2	153	Jatisari	2
124	Poncogati	3	154	Glingseran	2
125	Sumbersuko	3	155	Banyuputih	3
126	Silolembu	3	156	Wringin	2
127	Locare	2	157	Sumbercantin g	4
128	Sumber Salak	5	158	Karanganyar	1
129	Gadingsari	2	159	Sekarputih	3
130	Binakal	2	160	Tegalampel	3
131	Sumber Waru	2	161	Mandiro	5
132	Kembangan	2	162	Tanggulangun	2
133	Baratan	1	163	Klabang	2
134	Jeruk Soksok	2	164	Klabang Agung	1
135	Bendelan	2	165	Purnama	1
136	Sumber Tengah	2	166	Trebungan	1
137	Andungsari	1	167	Paguan	2
138	K U P A N G	2	168	Sumberkokap	2
139	Ardisaeng	2	169	Taman	1
140	Sumberdumpyo ng	1	170	Gentong	1
141	Gadingsari	2	171	Kemuningan	1
142	P A K E M	2	172	Kretek	2
143	Patemon	2	173	Karanganyar	1
144	P E T U N G	2	174	Blimbing	2
145	Ambulu	1	175	Karang Sengon	2
146	Bukor	2	176	Wonokerto	2
147	Sumbermalang	3	177	Klabang	1
148	Jambewungu	3	178	Besuk	2
149	Gubrih	2	179	Klampokan	2
150	Ampelan	2	180	Sumbersuko	3

Lampiran 13 Output Pengelompokan *Two Step Cluster* (Lanjutan)

No	Desa	Cluster	No	Desa	Cluster
181	Pandak	3	211	Ramban Wetan	3
182	Leprak	3	212	Ramban Kulon	3
183	Wonobojo	1	213	Grujugan	5
184	Sempol	2	214	Bajuran	1
185	Kalianyar	6	215	Jirek Mas	5
186	Kalisat	2	216	Batusalang	1
187	Jampit	6	217	Pelalangan	3
188	Sumber Rejo	6	218	Batu Ampar	6
189	Kaligedang	6	219	Bercak Asri	3
190	Sumbercanting	1			
191	Lanas	1			
192	Penang	6			
193	Klekean	2			
194	Gayam	5			
195	Lumutan	4			
196	Botolinggo	3			
197	Gayam Lor	3			
198	Tarum	3			
199	Bandilan	3			
200	Sempol	3			
201	Prajejan Lor	3			
202	Prajejan Kidul	3			
203	Cangkring	5			
204	Walidono	5			
205	Suling Kulon	3			
206	Solor	6			
207	Kladi	1			
208	Bercak	3			
209	Cermee	4			
210	Suling Wetan	3			

Lampiran 14 Syntax dan Output Mshapiro Test

```
> mshapiro.test(C)
```

Shapiro-Wilk normality test

data: Z

W = 0.62358, p-value < 2.2e-16

Lampiran 15 Output Pengujian Beda Mean MANOVA

	Effect	Value	F	Hypothesis df	Error df	Sig.
Cluster	Pillai's Trace	0.234	5.755b	11	207	0.000
	Wilks' Lambda	0.766	5.755b	11	207	0.000
	Hotelling's Trace	0.306	5.755b	11	207	0.000
	Roy's Largest Root	0.306	5.755b	11	207	0.000

Lampiran 16 Output Pengujian ANOVA**One-way ANOVA: X1 versus Cluster**

Source	DF	SS	MS	F	P
Cluster	1	1.428	1.428	6.88	0.009
Error	217	45.074	0.208		
Total	218	46.502			

One-way ANOVA: X2 versus Cluster

Source	DF	SS	MS	F	P
Cluster	1	8.609	8.609	20.22	0.000
Error	217	92.386	0.426		
Total	218	100.995			

One-way ANOVA: X3 versus Cluster

Source	DF	SS	MS	F	P
Cluster	1	7.80	7.80	6.70	0.010
Error	217	252.53	1.16		
Total	218	260.33			

One-way ANOVA: X4 versus Cluster

Source	DF	SS	MS	F	P
Cluster	1	0.9660	0.9660	11.63	0.001
Error	217	18.0203	0.0830		
Total	218	18.9863			

One-way ANOVA: X5 versus Cluster

Source	DF	SS	MS	F	P
Cluster	1	12228	12228	5.97	0.015
Error	217	444295	2047		
Total	218	456523			

Lampiran 16 Output Pengujian ANOVA (Lanjutan)**One-way ANOVA: X6 versus Cluster**

Source	DF	SS	MS	F	P
Cluster	1	4633	4633	21.18	0.000
Error	217	47478	219		
Total	218	52112			

One-way ANOVA: X7 versus Cluster

Source	DF	SS	MS	F	P
Cluster	1	20289	20289	5.39	0.021
Error	217	817107	3765		
Total	218	837397			

One-way ANOVA: X8 versus Cluster

Source	DF	SS	MS	F	P
Cluster	1	5280	5280	5.64	0.018
Error	217	203298	937		
Total	218	208577			

One-way ANOVA: X9 versus Cluster

Source	DF	SS	MS	F	P
Cluster	1	7.54	7.54	2.11	0.148
Error	217	777.18	3.58		
Total	218	784.72			

One-way ANOVA: X10 versus Cluster

Source	DF	SS	MS	F	P
Cluster	1	9.82	9.82	8.59	0.004
Error	217	248.11	1.14		
Total	218	257.93			

One-way ANOVA: X11 versus Cluster

Source	DF	SS	MS	F	P
Cluster	1	1839.2	1839.2	39.53	0.000
Error	217	10095.3	46.5		
Total	218	11934.6			

Lampiran 17 Output Uji Mann-Whitney**Mann-Whitney Test and CI: X12(1), X12(2)**

	N	Median
X12(1)	71	3.0000
X12(2)	148	5.0000

Point estimate for ETA1-ETA2 is -2.0000
 95.0 Percent CI for ETA1-ETA2 is (-2.0000,-2.0002)
 W = 4946.0
 Test of ETA1 = ETA2 vs ETA1 not = ETA2 is
 significant at 0.0000
 The test is significant at 0.0000 (adjusted for
 ties)

Mann-Whitney Test and CI: X13(1), X13(2)

	N	Median
X13(1)	71	1.0000
X13(2)	148	5.0000

Point estimate for ETA1-ETA2 is -4.0000
 95.0 Percent CI for ETA1-ETA2 is (-4.0002,-3.9998)
 W = 4023.5
 Test of ETA1 = ETA2 vs ETA1 not = ETA2 is
 significant at 0.0000
 The test is significant at 0.0000 (adjusted for
 ties)

Mann-Whitney Test and CI: X14(1), X14(2)

	N	Median
X14(1)	71	2.0000
X14(2)	148	3.0000

Point estimate for ETA1-ETA2 is -1.0000
 95.0 Percent CI for ETA1-ETA2 is (-1.0000,-1.0000)
 W = 4454.0
 Test of ETA1 = ETA2 vs ETA1 not = ETA2 is
 significant at 0.0000
 The test is significant at 0.0000 (adjusted for
 ties)

Lampiran 17 Output Uji Mann-Whitney (Lanjutan)**Mann-Whitney Test and CI: X15(1), X15(2)**

	N	Median
X15(1)	71	2.0000
X15(2)	148	3.0000

Point estimate for ETA1-ETA2 is -1.0000
 95.0 Percent CI for ETA1-ETA2 is (-1.0001,-1.0000)
 W = 5407.5
 Test of ETA1 = ETA2 vs ETA1 not = ETA2 is significant
 at 0.0000
 The test is significant at 0.0000 (adjusted for ties)

Mann-Whitney Test and CI: X16(1), X16(2)

	N	Median
X16(1)	71	0.0000
X16(2)	148	2.0000

Point estimate for ETA1-ETA2 is 0.0000
 95.0 Percent CI for ETA1-ETA2 is (-2.0000,-0.0002)
 W = 6456.5
 Test of ETA1 = ETA2 vs ETA1 not = ETA2 is significant
 at 0.0021
 The test is significant at 0.0008 (adjusted for ties)

Mann-Whitney Test and CI: X17(1), X17(2)

	N	Median
X17(1)	71	4.0000
X17(2)	148	4.0000

Point estimate for ETA1-ETA2 is -1.0000
 95.0 Percent CI for ETA1-ETA2 is (-1.0000,-1.0001)
 W = 5374.5
 Test of ETA1 = ETA2 vs ETA1 not = ETA2 is significant
 at 0.0000
 The test is significant at 0.0000 (adjusted for ties)

Lampiran 17 Output Uji Mann-Whitney (Lanjutan)**Mann-Whitney Test and CI: X18(1), X18(2)**

	N	Median
X18(1)	71	5.0000
X18(2)	148	5.0000

Point estimate for ETA1-ETA2 is -0.0000
 95.0 Percent CI for ETA1-ETA2 is (0.0000,-0.0000)
 W = 7771.0
 Test of ETA1 = ETA2 vs ETA1 not = ETA2 is significant
 at 0.9301
 The test is significant at 0.5945 (adjusted for ties)

Mann-Whitney Test and CI: X19(1), X19(2)

	N	Median
X19(1)	71	3.0000
X19(2)	148	3.0000

Point estimate for ETA1-ETA2 is -0.0000
 95.0 Percent CI for ETA1-ETA2 is (0.0000,-0.0001)
 W = 7018.0
 Test of ETA1 = ETA2 vs ETA1 not = ETA2 is significant
 at 0.0713
 The test is significant at 0.0223 (adjusted for ties)

Mann-Whitney Test and CI: X20(1), X20(2)

	N	Median
X20(1)	71	3.000
X20(2)	148	4.000

Point estimate for ETA1-ETA2 is -0.000
 95.0 Percent CI for ETA1-ETA2 is (-1.000,0.000)
 W = 6536.5
 Test of ETA1 = ETA2 vs ETA1 not = ETA2 is significant
 at 0.0037
 The test is significant at 0.0020 (adjusted for ties)

Lampiran 17 Output Uji Mann-Whitney (Lanjutan)**Mann-Whitney Test and CI: X21(1), X21(2)**

	N	Median
X21(1)	71	4.0000
X21(2)	148	4.0000

Point estimate for ETA1-ETA2 is -0.0000

95.0 Percent CI for ETA1-ETA2 is (0.0003,-0.0002)

W = 7107.0

Test of ETA1 = ETA2 vs ETA1 not = ETA2 is significant
at 0.1095

The test is significant at 0.0639 (adjusted for ties)

Lampiran 18 Surat Keterangan Data dari BPS



SURAT KETERANGAN

Yang bertanda tangan dibawah ini :

N a m a : Thomas Wunang Tjahjo, M.Sc, M.Eng.
 N I P : 19700329 1992 11 1 001
 Jabatan : Kepala Bidang Integrasi Pengolahan dan
 Diseminasi Statistik

Dengan ini menerangkan bahwa :

N a m a : Rahmat Ramdhany
 Fakultas/Program Studi : Matematika Dan Ilmu Pengetahuan Alam / Statistika
 N.R.P : 06211440000004
 Alamat Rumah : Perumahan Dosen ITS Blok D Jalan Tekbik Industri No 29
 Surabaya
 Akademi / Universitas : Institut Teknologi Sepuluh Nopember (ITS)
 Telp (031) 594 3352, (031) 599 4251-55
 Fax (031) 592 2940

Di berikan kesempatan menggunakan data Badan Pusat Statistik (BPS) Provinsi Jawa Timur, dengan syarat menyebut judul publikasi dan sumbernya serta tidak untuk tujuan komersil. Data ini digunakan dalam rangka menyusun Tugas Akhir / Skripsi / Thesis / Disertasi dengan judul :

"Pengelompokan Desa di Kabupaten Bondowoso berdasarkan Data Campuran Numerik dan Kategori menggunakan metode Ensemble Rock "

Demikian surat keterangan ini dibuat dan agar dipergunakan sebagaimana mestinya

Surabaya, 4 Januari 2018

An. Kepala BPS Provinsi Jawa Timur

Kepala Bidang IPDS

Thomas Wunang Tjahjo, M.Sc, M.Eng.



(Halaman ini sengaja dikosongkan)

BIODATA PENULIS



Penulis memiliki nama lengkap Rahmat Ramdhany yang biasa dipanggil Dhany. Penulis lahir di Mataram pada tanggal 28 Januari 1996. Penulis merupakan anak ketiga dari Bapak H. Hurrani dan Ibu Hj. Siti Rochaini. Penulis menempuh pendidikan SD di SDN 16 Mataram, pendidikan SMP di SMP Negeri 2 Mataram dan SMA di SMA Negeri 1 Mataram. Hingga akhirnya pada tahun 2014 penulis melanjutkan pendidikan di jenjang perguruan tinggi di departemen Statistika ITS melalui jalur SNMPTN Undangan yang terdaftar dengan NRP. 1314100004. Selama 3.5 tahun berkuliah di departemen Statistika ITS, penulis juga aktif di Himpunan Mahasiswa Statitika ITS (HIMASTA-ITS) sebagai staff Departemen Minat Bakat dan delegasi HIMASTA-ITS dalam Ikatan Himpunan Mahasiswa Seluruh Indonesia (IHMSI). Penulis juga berkesempatan menjadi Asisten Dosen untuk matakuliah Pengantar Metode Statistika Segala kritik dan saran serta diskusi lebih lanjut mengenai Tugas Akhir ini dapat dikirimkan melalui surat elektronik (*e-mail*) ke alamat ramdhanyrahmat@gmail.com.